

JFW



Attorney Docket No.: 2543-1-033

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

APPLICANTS : Jonathan A. Terrett  
SERIAL NO. : 10/736,227  
FILED : December 15, 2003  
FOR : NOVEL CANCER ASSOCIATED PROTEIN

**Certificate of Mailing Under 37 CFR 1.8**

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to COMMISSIONER FOR PATENTS, P.O. Box 1450, Alexandria, VA 22313-1450 on May 17, 2007.

Carolyn Di Meglio  
(Name of Depositor)

 5/17/07  
(Signature and Date)

**PETITION FOR GRANT OF PRIORITY UNDER 35 USC 119**

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Dear Sir:

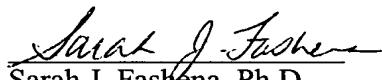
Applicant hereby petitions for grant of priority of the present Application on the basis of the following prior filed foreign Application:

<u>COUNTRY</u>	<u>SERIAL NO.</u>	<u>FILING DATE</u>
Great Britain	GB0205264.5	MARCH 6, 2002

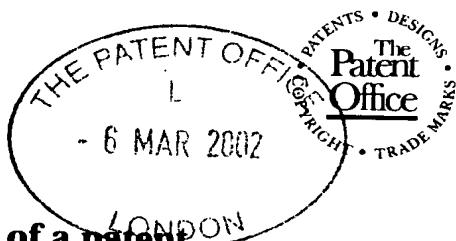
To perfect Applicant's claim to priority, a certified copy of the above listed prior filed Application is enclosed.

Acknowledgment of Applicant's perfection of claim to priority is accordingly requested.

Respectfully submitted,

  
Sarah J. Fashena, Ph.D.  
Agent for Applicant  
Registration No. 57,600

KLAUBER & JACKSON LLC  
411 Hackensack Avenue  
Hackensack, NJ 07601  
(201)487-5800  
Date: May 17, 2007

07MAR02 E701271-1 D10103  
F01/7700 0.00-0205264.5

The Patent Office

Cardiff Road  
Newport  
South Wales  
NP10 8QQ

## Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

1. Your reference P0081-GB02

2. Patent application number  
(The Patent Office will fill in this part)

0205264.5

3. Full name, address and postcode of the or of each applicant (underline all surnames)

OXFORD GLYCOSCIENCES (UK) LTD  
THE FORUM  
86 MILTON PARK  
ABINGDON, OXON OX14 4RY  
UNITED KINGDOM

Patents ADP number (if you know it)

7112386002

If the applicant is a corporate body, give the country/state of its incorporation

ENGLAND AND WALES

4. Title of the invention

PROTEIN

5. Name of your agent (if you have one)

BLAKEY, ALISON JANE  
OXFORD GLYCOSCIENCES (UK) LTD  
THE FORUM  
86 MILTON PARK  
ABINGDON, OXON OX14 4RY  
UNITED KINGDOM

Dr John Thompson  
UCB Celltech  
208 Bath Rd  
Slough  
Berkshire  
SL1 3WE

5 1 77  
19/4b  
Pm.  
Patents ADP number (if you know it)

7112386002

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (if you know it) the or each application number

Country

Priority application number  
(if you know it)Date of filing  
(day / month / year)

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application

Date of filing  
(day / month / year)

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? (Answer 'Yes' if:

YES

- a) any applicant named in part 3 is not an inventor, or
- b) there is an inventor who is not named as an applicant, or
- c) any named applicant is a corporate body.

See note (d))

# Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form.  
Do not count copies of the same document

Continuation sheets of this form -

Description 33

Claim(s) 3

Abstract 1

Drawing(s) 247

---

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (Patents Form 7/77)

Request for preliminary examination and search (Patents Form 9/77)

Request for substantive examination  
(Patents Form 10/77)

Any other documents  
(please specify)

---

11.

I/We request the grant of a patent on the basis of this application.

Signature Allison Blaney Date 6/03/02

---

12. Name and daytime telephone number of person to contact in the United Kingdom

MARY GADSDEN 01235 208127

## Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

## Notes

- a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.
- b) Write your answers in capital letters using black ink or you may type them.
- c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- e) Once you have filled in the form you must remember to sign and date it.
- f) For details of the fee and ways to pay please contact the Patent Office.

## PROTEIN

5 The present invention relates to a new protein isolated from breast cancer cell line membrane preparations, compositions comprising the protein, including vaccines and antibodies which are immunospecific for the protein.

10 Breast cancer is the most frequently diagnosed cancer in women. The implementation of screening programs for the early detection of breast cancer, and the advent of anticancer treatments, such as chemotherapy, radiotherapy and anti-oestrogen therapies, to augment surgical resection have improved the survival of breast cancer patients.

15 However, some breast tumours become refractory to such treatments, as the cancer cells develop resistance to chemotherapy drugs or lose their hormone sensitivity, leading to recurrent or metastatic disease which is often incurable. More recently, attention has focussed on the development of immunological therapies (Green, M.C., *et al.* *Cancer Treat. Rev.* **26**, 269-286 (2000); Davis, I.D., *et al.* *Immunol. Cell Biol.* **78**, 179-195 (2000); Knuth, A., *et al.* *Cancer Chemother Pharmacol.* **46**, S46-51 (2000); Shiku, H., *et al.* *Cancer Chemother. Pharmacol.* **46**, S77-82 (2000); Saffran, D.C., *et al.* *Cancer Metastasis Rev.* **18**, 437-449 (1999)), such as cancer vaccines and monoclonal antibodies (mAbs), as a means of initiating 20 and targeting a host immune response against tumour cells. In 1998 the FDA approved the use of Herceptin (Stebbing, J., *et al.* *Cancer Treat. Rev.* **26**, 287-290 (2000); Dillman, R.O. *Cancer Biother. Radiopharm.* **14**, 5-10 (1999); Miller, K.D., *et al.* *Invest. New Drugs* **17**, 417-427 (1999)), a mAb that recognises the erbB2/HER2-neu receptor protein, as a treatment for metastatic breast cancer. In combination with chemotherapy, Herceptin has been shown to 25 prolong the time to disease progression, when compared to patients receiving chemotherapy alone (Baselga, J., *et al.* *Cancer Res.* **58**, 2825-2831 (1998)). Herceptin, however, is only effective in treating the 10-20% of patients whose tumours over-express the erbB2 protein. Thus, the identification of other suitable targets or antigens for immunotherapy of breast 30 cancer has become increasingly important.

35 Kidney cancer is diagnosed in more than 28,000 patients each year in the United States (<http://cancernet.nci.nih.gov>). Renal Cell Cancer is the most common form of kidney cancer in adults. Transitional Cell Cancer (carcinoma), which affects the renal pelvis, is a less common form of kidney cancer. It is similar to cancer that occurs in the bladder and is often treated like bladder cancer. Finally, Wilms' tumor, the most common type of childhood kidney cancer, is different from kidney cancer in adults.

40 Kidney cancer is usually treated with surgery, radiation therapy, and/or biological therapy (Interleukin-2 and Interferon-alpha), and, to a lesser extent, chemotherapy and hormone therapy.

An ideal protein target for cancer immunotherapy should have a restricted expression profile in normal tissues and be over-expressed in tumours, such that the immune response will be targeted to tumour cells and not against other organs. In addition, the protein target should be exposed on the cell surface, where it will be accessible to therapeutic agents. Tumour antigens

have been identified for a number of cancer types, by using techniques such as differential screening of cDNA (Hubert, R.S., *et al.* *Proc. Natl. Acad. Sci. USA* **96**, 14523-14528 (1999); Lucas, S., *et al.* *Int. J. Cancer* **87**, 55-60 (2000)), and the purification of cell-surface antigens that are recognised by tumour-specific antibodies (Catimel, B., *et al.* *J. Biol. Chem.* **271**, 25664-25670 (1996)). As an alternative approach to identifying breast tumour antigens, we have used proteomics to characterise the complement of proteins in cell membranes isolated from the breast cancer cell line MDA-MB-468. In this way, we have identified a protein, designated BCMP 101, which shows restricted expression to a few tissues, with elevated expression in breast tumours and kidney cancer cell lines, suggesting that it may be a suitable target for cancer therapy and diagnosis.

Thus, in a first aspect, the present invention provides a polypeptide which:

- a) comprises or consists of the amino acid sequence shown in Figure 1 (SEQ ID NO: 1);
- 15 b) is a derivative having one or more amino acid substitutions, deletions or insertions relative to the amino acid sequence shown in Figure 1 (SEQ ID NO: 1); or
- c) is a fragment of a polypeptide as defined in a) or b) above, which is at least ten amino acids long.

20 Polypeptides of the present invention may be in substantially pure, isolated or recombinant form, and may be fused to other moieties. In particular, fusions of the polypeptides of the present invention with localisation-reporter proteins such as the Green Fluorescent Protein (U.S. Patent Nos. 5,625,048, 5,777,079, 6,054,321 and 5,804,387) or the DsRed fluorescent protein 25 (Matz, M. V., *et al.* *Nature Biotech.* **17**:969-973.) are specifically contemplated by the present invention. They are provided in substantially pure form; that is to say, they are free, to a substantial extent, from other proteins. Thus, a polypeptide of the present invention may be provided in a composition in which it is the predominant component present (i.e. it is present at a level of at least 50%; preferably at least 75%, at least 90%, or at least 95%; when determined on 30 a weight/weight basis excluding solvents or carriers).

In order to more fully appreciate the present invention, polypeptides within the scope of a)-c) above will now be discussed in greater detail.

35 Polypeptides within the scope of a)

A polypeptide within the scope of a), may consist of the particular amino acid sequence given in Figure 1 (SEQ ID NO:1) or may have an additional N-terminal and/or an additional C-terminal amino acid sequence relative to the sequence given in Figure 1 (SEQ ID NO: 1).

40 Additional N-terminal or C-terminal sequences may be provided for various reasons. Techniques for providing such additional sequences are well known in the art.

Additional sequences may be provided in order to alter the characteristics of a particular polypeptide. This can be useful in improving expression or regulation of expression in particular expression systems. For example, an additional sequence may provide some protection against proteolytic cleavage. This has been done for the hormone Somatostatin by fusing it at its N-terminus to part of the  $\beta$  galactosidase enzyme (Itakwa *et al.*, *Science* **198**: 105-63 (1977)).

5

Additional sequences can also be useful in altering the properties of a polypeptide to aid in identification or purification. For example, a fusion protein may be provided in which a polypeptide is linked to a moiety capable of being isolated by affinity chromatography. The 10 moiety may be an antigen or an epitope and the affinity column may comprise immobilised antibodies or immobilised antibody fragments which bind to said antigen or epitope (desirably with a high degree of specificity). The fusion protein can usually be eluted from the column by addition of an appropriate buffer.

15 Additional N-terminal or C-terminal sequences may, however, be present simply as a result of a particular technique used to obtain a polypeptide of the present invention and need not provide any particular advantageous characteristic to the polypeptide of the present invention. Such polypeptides are within the scope of the present invention.

20 Whatever additional N-terminal or C-terminal sequence is present, it is preferred that the resultant polypeptide should exhibit the immunological activity of the polypeptide having the amino acid sequence shown in Figure 1 (SEQ ID NO: 1).

Polypeptides within the scope of b)

25

Turning now to the polypeptides defined in b) above, it will be appreciated by the person skilled in the art that these polypeptides are variants of the polypeptide given in a) above, provided that such variants exhibit the immunological activity of the polypeptide having the amino acid sequence shown in Figure 1 (SEQ ID NO: 1).

30

Alterations in the amino acid sequence of a protein can occur which do not affect the function of a protein. These include amino acid deletions, insertions and substitutions and can result from alternative splicing and/or the presence of multiple translation start sites and stop sites. Polymorphisms may arise as a result of the infidelity of the translation process. Thus changes 35 in amino acid sequence may be tolerated which do not affect the protein's function.

The skilled person will appreciate that various changes can often be made to the amino acid sequence of a polypeptide which has a particular activity to produce variants (sometimes known as "muteins") having at least a proportion of said activity, and preferably having a substantial 40 proportion of said activity. Such variants of the polypeptides described in a) above are within the scope of the present invention and are discussed in greater detail below. They include allelic and non-allelic variants.

An example of a variant of the present invention is a polypeptide as defined in a) above, apart from the substitution of one or more amino acids with one or more other amino acids. The skilled person is aware that various amino acids have similar properties. One or more such amino acids of a substance can often be substituted by one or more other such amino acids without eliminating a desired activity of that substance.

Thus, the amino acids glycine, alanine, valine, leucine and isoleucine can often be substituted for one another (amino acids having aliphatic side chains). Of these possible substitutions, it is preferred that glycine and alanine are used to substitute for one another (since they have relatively short side chains) and that valine, leucine and isoleucine are used to substitute for one another (since they have larger aliphatic side chains which are hydrophobic).

Other amino acids which can often be substituted for one another include:

- phenylalanine, tyrosine and tryptophan (amino acids having aromatic side chains);
- lysine, arginine and histidine (amino acids having basic side chains);
- aspartate and glutamate (amino acids having acidic side chains);
- asparagine and glutamine (amino acids having amide side chains); and
- cysteine and methionine (amino acids having sulphur-containing side chains).

Substitutions of this nature are often referred to as "conservative" or "semi-conservative" amino acid substitutions.

Amino acid deletions or insertions may also be made relative to the amino acid sequence given in a) above. Thus, for example, amino acids which do not have a substantial effect on the activity of the polypeptide, or at least which do not eliminate such activity, may be deleted. Such deletions can be advantageous since the overall length and the molecular weight of a polypeptide can be reduced whilst still retaining activity. This can enable the amount of polypeptide required for a particular purpose to be reduced - for example, dosage levels can be reduced.

Amino acid insertions relative to the sequence given in a) above can also be made. This may be done to alter the properties of a polypeptide of the present invention (e.g. to assist in identification, purification or expression, as explained above in relation to fusion proteins).

Amino acid changes relative to the sequence given in a) above can be made using any suitable technique e.g. by using site-directed mutagenesis (Hutchinson et al., 1978, J. Biol. Chem. 253:6551).

It should be appreciated that amino acid substitutions or insertions within the scope of the present invention can be made using naturally occurring or non-naturally occurring amino acids. Whether or not natural or synthetic amino acids are used, it is preferred that only L- amino acids are present.

Whatever amino acid changes are made (whether by means of substitution, insertion or deletion), preferred polypeptides of the present invention have at least 50% sequence identity

with a polypeptide as defined in a) above, more preferably the degree of sequence identity is at least 75%. Sequence identities of at least 90% or at least 95% are most preferred.

The term identity can be used to describe the similarity between two polypeptide sequences.

5 The degree of amino acid sequence identity can be calculated using a program such as "bestfit" (Smith and Waterman, *Advances in Applied Mathematics*, 482-489 (1981)) to find the best segment of similarity between any two sequences. The alignment is based on 10 maximising the score achieved using a matrix of amino acid similarities, such as that described by Schwarz and Dayhof (1979) *Atlas of Protein Sequence and Structure*, Dayhof, M.O., Ed pp 353-358.

A software package well known in the art for carrying out this procedure is the CLUSTAL program. It compares the amino acid sequences of two polypeptides and finds the optimal 15 alignment by inserting spaces in either sequence as appropriate. The amino acid identity or similarity (identity plus conservation of amino acid type) for an optimal alignment can also be calculated using a software package such as BLASTx. This program aligns the largest stretch of similar sequence and assigns a value to the fit. For any one pattern comparison, several 20 regions of similarity may be found, each having a different score. One skilled in the art will appreciate that two polypeptides of different lengths may be compared over the entire length of the longer fragment. Alternatively small regions may be compared. Normally sequences of 25 the same length are compared for a useful comparison to be made.

Where high degrees of sequence identity are present there will be relatively few differences in amino acid sequence. Thus for example they may be less than 20, less than 10, or even less than 5 differences.

#### Polypeptides within the scope of c)

As discussed *supra*, it is often advantageous to reduce the length of a polypeptide, provided that 30 the resultant reduced length polypeptide still has a desired activity or can give rise to useful antibodies. Feature c) of the present invention therefore covers fragments of polypeptides a) or b) above.

The skilled person can determine whether or not a particular fragment has activity using the 35 techniques disclosed above. Preferred fragments are at least 10 amino acids long. They may be at least 20, at least 50 or at least 100 amino acids long.

As will be discussed below, the polypeptides of the present invention will find use in an immunotherapeutic approach to breast and/or kidney cancer. The skilled person will 40 appreciate that for the preparation of one or more polypeptides of the invention, the preferred approach will be based on recombinant DNA techniques. Thus, in a second aspect, the present invention provides an isolated or recombinant nucleic acid molecule which:

5

- a) comprises or consists of the DNA sequence shown in Figure 1 (SEQ ID NO:2) or its RNA equivalent;
- b) a sequence which is complementary to the sequences of a);
- c) a sequence which codes for the same or polypeptide, as the sequences of a) or b);
- d) a sequence which shows substantial identity with any of those of a), b) and c); or
- e) a sequence which codes for a derivative or fragment of a polypeptide shown in Figure 1 (SEQ ID NO:2).

10 These nucleic acid molecules are now discussed in greater detail.

15 The term identity can also be used to describe the similarity between two individual DNA sequences. The 'bestfit' program (Smith and Waterman, *Advances in applied Mathematics*, 482-489 (1981)) is one example of a type of computer software used to find the best segment of similarity between two nucleic acid sequences, whilst the GAP program enables sequences to be aligned along their whole length and finds the optimal alignment by inserting spaces in either sequence as appropriate. It is preferred if sequences which show substantial identity with any of those of a), b) and c) have e.g. at least 50%, at least 75% or at least 90% or 95% sequence identity.

20 The polypeptides of the present invention can be coded for by a large variety of nucleic acid molecules, taking into account the well known degeneracy of the genetic code. All of these molecules are within the scope of the present invention. They can be inserted into vectors and cloned to provide large amounts of DNA or RNA for further study. Suitable vectors may be 25 introduced into host cells to enable the expression of polypeptides of the present invention using techniques known to the person skilled in the art.

30 The term 'RNA equivalent' when used above indicates that a given RNA molecule has a sequence which is complementary to that of a given DNA molecule, allowing for the fact that in RNA 'U' replaces 'T' in the genetic code. The nucleic acid molecule may be in isolated, recombinant or chemically synthetic form.

35 Techniques for cloning, expressing and purifying proteins and polypeptides are well known to the skilled person. DNA constructs can readily be generated using methods well known in the art. These techniques are disclosed, for example in J. Sambrook *et al*, *Molecular Cloning 2<sup>nd</sup> Edition*, Cold Spring Harbour Laboratory Press (1989); in Old & Primrose *Principles of Gene Manipulation* 5th Edition, Blackwell Scientific Publications (1994); and in Stryer [*Biochemistry* 4th Edition, W H Freeman and Company (1995)]. Modifications of DNA constructs and the 40 proteins expressed such as the addition of promoters, enhancers, signal sequences, leader sequences, translation start and stop signals and DNA stability controlling regions, or the addition of fusion partners may then be facilitated.

Normally the DNA construct will be inserted into a vector, which may be of phage or plasmid origin. Expression of the protein is achieved by the transformation or transfection of the

vector into a host cell which may be of eukaryotic or prokaryotic origin. Such vectors and suitable host cells form third and fourth aspects of the present invention.

Knowledge of the nucleic acid structure can be used to raise antibodies and for gene therapy.

5 Techniques for this are well-known by those skilled in the art.

By using appropriate expression systems, polypeptides of the present invention may be expressed in glycosylated or non-glycosylated form. Non-glycosylated forms can be produced by expression in prokaryotic hosts, such as *E. coli*.

10

Polypeptides comprising N-terminal methionine may be produced using certain expression systems, whilst in others the mature polypeptide will lack this residue.

Preferred techniques for cloning, expressing and purifying a substance of the present invention are summarised below:

15

Polypeptides may be prepared natively or under denaturing conditions and then subsequently refolded. Baculoviral expression vectors include secretory plasmids (such as pACGP67 from Pharmingen), which may have an epitope tag sequence cloned in frame (e.g. myc, V5 or His) to aid detection and allow for subsequent purification of the protein. Mammalian expression

20

vectors may include pCDNA3 and pSecTag (both Invitrogen), and pREP9 and pCEP4 (invitrogen). *E. coli* systems include the pBad series (His tagged - Invitrogen) or pGex series (Pharmacia).

25

In addition to nucleic acid molecules coding for polypeptides according to the present invention, referred to herein as "coding" nucleic acid molecules, the present invention also includes nucleic acid molecules complementary thereto. Thus, for example, both strands of a double stranded nucleic acid molecule are included within the scope of the present invention (whether or not they are associated with one another). Also included are mRNA molecules and complementary DNA Molecules (e.g. cDNA molecules).

30

Nucleic acid molecules which can hybridise to any of the nucleic acid molecules discussed above are also covered by the present invention. Such nucleic acid molecules are referred to herein as "hybridising" nucleic acid molecules. Hybridising nucleic acid molecules can be useful as probes or primers, for example.

35

Desirably such hybridising molecules are at least 10 nucleotides in length and preferably are at least 25 or at least 50 nucleotides in length. The hybridising nucleic acid molecules preferably hybridise to nucleic acids within the scope of (a), (b), (c), (d) or (e) above specifically.

40

Desirably the hybridising molecules will hybridise to such molecules under stringent hybridisation conditions. One example of stringent hybridisation conditions is where attempted hybridisation is carried out at a temperature of from about 35°C to about 65°C using a salt solution which is about 0.9 molar. However, the skilled person will be able to vary such

conditions as appropriate in order to take into account variables such as probe length, base composition, type of ions present, etc.

5 Manipulation of the DNA encoding the protein is a particularly powerful technique for both modifying proteins and for generating large quantities of protein for purification purposes.

This may involve the use of PCR techniques to amplify a desired nucleic acid sequence. Thus the sequence data provided herein can be used to design primers for use in PCR so that a desired sequence can be targeted and then amplified to a high degree.

10 Typically, primers will be at least five nucleotides long and will generally be at least ten nucleotides long (e.g. fifteen to twenty-five nucleotides long). In some cases, primers of at least thirty or at least thirty-five nucleotides in length may be used.

15 As a further alternative chemical synthesis may be used. This may be automated. Relatively short sequences may be chemically synthesised and ligated together to provide a longer sequence.

20 In addition to being used as primers and/or probes, hybridising nucleic acid molecules of the present invention can be used as anti-sense molecules to alter the expression of substances of the present invention by binding to complementary nucleic acid molecules. This technique can be used in anti-sense therapy.

25 A hybridising nucleic acid molecule of the present invention may have a high degree of sequence identity along its length with a nucleic acid molecule within the scope of (a)-(e) above (e.g. at least 50%, at least 75% or at least 90% or 95% sequence identity). As will be appreciated by the skilled person, the higher the sequence identity a given single stranded nucleic acid molecule has with another nucleic acid molecule, the greater the likelihood that it will hybridise to a nucleic acid molecule which is complementary to that other nucleic acid molecule under appropriate conditions.

30 In view of the foregoing description the skilled person will appreciate that a large number of nucleic acids are within the scope of the present invention. Unless the context indicates otherwise, nucleic acid molecules of the present invention may have one or more of the following characteristics:

35

- 1) they may be DNA or RNA;
- 2) they may be single or double stranded;
- 3) they may be provided in recombinant form i.e. covalently linked to a 5' and/or a 3' flanking sequence to provide a molecule which does not occur in nature;
- 40 4) they may be provided without 5' and/or 3' flanking sequences which normally occur in nature;
- 5) they may be provided in substantially pure form. Thus they may be provided in a form which is substantially free from contaminating proteins and/or from other nucleic acids; and

6) they may be provided with introns or without introns (e.g. as cDNA).

As described herein, BCMP 101 is associated with breast and kidney cancer and as such provides a means of detection/diagnosis. Thus, in a fifth aspect, the present invention provides a method of screening for and/or diagnosis of breast and/or kidney cancer in a subject which comprises the step of detecting and/or quantifying the amount of a polypeptide of the invention in a biological sample obtained from said subject.

10 A convenient means for such detection/quantifying will involve the use of antibodies. Thus, the polypeptides of the invention also find use in raising antibodies. Thus, in a sixth aspect, the present invention provides antibodies, which bind to a polypeptide of the present invention or to a fragment of such a polypeptide. Preferred antibodies bind specifically to polypeptides of the present invention so that they can be used to purify and/or inhibit the activity of such polypeptides. The antibodies may be monoclonal or polyclonal.

15 Thus, the polypeptide of the invention, its fragments or other derivatives, or analogs thereof, may be used as an immunogen to generate antibodies which immunospecifically bind such an immunogen. Antibodies of the invention include, but are not limited to polyclonal, monoclonal, bispecific, humanized or chimeric antibodies, single chain antibodies, Fab fragments and F(ab') fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that specifically binds an antigen. The immunoglobulin molecules of the invention can be 20 of any class (e.g., IgG, IgE, IgM, IgD and IgA) or subclass of immunoglobulin molecule.

25 In the production of antibodies, screening for the desired antibody can be accomplished by techniques known in the art, e.g. ELISA (enzyme-linked immunosorbent assay). For example, to select antibodies which recognize a specific domain of a polypeptide of the invention, one may assay generated hybridomas for a product which binds to a polypeptide fragment containing such domain. For selection of an antibody that specifically binds a first polypeptide homolog but which does not specifically bind to (or binds less avidly to) a second polypeptide homolog, one can select on the basis of positive binding to the first polypeptide homolog and a lack of binding to (or reduced binding to) the second polypeptide homolog.

35 For preparation of monoclonal antibodies (mAbs) directed toward a polypeptide of the invention or fragment or analog thereof, any technique which provides for the production of antibody molecules by continuous cell lines in culture may be used. For example, the hybridoma technique originally developed by Kohler and Milstein (1975, *Nature* 256:495-497), as well as the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 40 1983, *Immunology Today* 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., 1985, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the

mAbs of the invention may be cultivated in vitro or in vivo. In an additional embodiment of the invention, mAbs can be produced in germ-free animals utilizing known technology (PCT/US90/02545, incorporated herein by reference).

5 The mAbs include but are not limited to human mAbs and chimeric mAbs (e.g., human-mouse chimeras). A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a human immunoglobulin constant region and a variable region derived from a murine mAb. (See, e.g., U.S. 4,816,567; and U.S. 4,816,397, which are incorporated herein by reference in their entirety.) Humanized

10 antibodies are antibody molecules from non-human species having one or more complementarity determining regions (CDRs) from the non-human species and a framework region from a human immunoglobulin molecule. (See, e.g., U.S. 5,585,089, which is incorporated herein by reference in its entirety).

15 Chimeric and humanized mAbs can be produced by recombinant DNA techniques known in the art, for example using methods described in WO 87/02671; EP 184,187; EP 171,496; EP 173,494; WO 86/01533; U.S. 4,816,567; EP 125,023; Better et al., 1988, *Science* 240:1041-1043; Liu et al., 1987, *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu et al., 1987, *J. Immunol.* 139:3521-3526; Sun et al., 1987, *Proc. Natl. Acad. Sci. USA* 84:214-218;

20 Nishimura et al., 1987, *Canc. Res.* 47:999-1005; Wood et al., 1985, *Nature* 314:446-449; and Shaw et al., 1988, *J. Natl. Cancer Inst.* 80:1553-1559; Morrison, 1985, *Science* 229:1202-1207; Oi et al., 1986, *Bio/Techniques* 4:214; U.S. 5,225,539; Jones et al., 1986, *Nature* 321:552-525; Verhoeven et al. (1988) *Science* 239:1534; and Beidler et al., 1988, *J. Immunol.* 141:4053-4060.

25 Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Such antibodies can be produced using transgenic mice which are incapable of expressing endogenous immunoglobulin heavy and light chain genes, but which can express human heavy and light chain genes. The transgenic mice are immunized in the normal

30 fashion with a selected antigen, e.g., all or a portion of a polypeptide of the invention. mAbs directed against the antigen can be obtained using conventional hybridoma technology. The human immunoglobulin transgenes harboured by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA, IgM and IgE

35 antibodies. For an overview of this technology for producing human antibodies, see Lonberg and Huszar (1995, *Int. Rev. Immunol.* 13:65-93). For a detailed discussion of this technology for producing human antibodies and human mAbs and protocols for producing such

40 antibodies, see, e.g., U.S. 5,625,126; U.S. 5,633,425; U.S. 5,569,825; U.S. 5,661,016; and U.S. 5,545,806. In addition, companies such as Abgenix, Inc. (Freemont, CA) and Genpharm (San Jose, CA) can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human mAb, e.g.,

a mouse antibody, is used to guide the selection of a completely human antibody recognizing the same epitope. (Jespers et al. (1994) *Bio/technology* 12:899-903).

5 The antibodies of the present invention can also be generated using various phage display methods known in the art. In phage display methods, functional antibody domains are displayed on the surface of phage particles which carry the polynucleotide sequences encoding them. In a particular, such phage can be utilized to display antigen-binding domains expressed from a repertoire or combinatorial antibody library (e.g., human or murine). Phage expressing an antigen binding domain that binds the antigen of interest can be selected or  
10 identified with antigen, e.g., using labelled antigen or antigen bound or captured to a solid surface or bead. Phage used in these methods are typically filamentous phage including fd and M13 binding domains expressed from phage with Fab, Fv or disulfide stabilized Fv antibody domains recombinantly fused to either the phage gene III or gene VIII protein. Phage display methods that can be used to make the antibodies of the present invention  
15 include those disclosed in Brinkman et al., *J. Immunol. Methods* 182:41-50 (1995); Ames et al., *J. Immunol. Methods* 184:177-186 (1995); Kettleborough et al., *Eur. J. Immunol.* 24:952-958 (1994); Persic et al., *Gene* 187 9-18 (1997); Burton et al., *Advances in Immunology* 57:191-280 (1994); PCT Application No. PCT/GB91/01134; WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; and U.S. Patent  
20 Nos. 5,698,426; 5,223,409; 5,403,484; 5,580,717; 5,427,908; 5,750,753; 5,821,047; 5,571,698; 5,427,908; 5,516,637; 5,780,225; 5,658,727; 5,733,743 and 5,969,108; each of which is incorporated herein by reference in its entirety.

25 As described in the above references, after phage selection, the antibody coding regions from the phage can be isolated and used to generate whole antibodies, including human antibodies, or any other desired antigen binding fragment, and expressed in any desired host, including mammalian cells, insect cells, plant cells, yeast, and bacteria, e.g., as described in detail below. For example, techniques to recombinantly produce Fab, Fab' and F(ab')<sup>2</sup> fragments can also be employed using methods known in the art such as those disclosed in WO  
30 92/22324; Mullinax et al., *BioTechniques* 12(6):864-869 (1992); and Sawai et al., *AJRI* 34:26-34 (1995); and Better et al., *Science* 240:1041-1043 (1988) (said references incorporated by reference in their entireties).

35 Examples of techniques which can be used to produce single-chain Fvs and antibodies include those described in U.S. 4,946,778 and U.S. 5,258,498; Huston et al., *Methods in Enzymology* 203:46-88 (1991); Shu et al., *PNAS* 90:7995-7999 (1993); and Skerra et al., *Science* 240:1038-1040 (1988).

40 The invention further provides for the use of bispecific antibodies, which can be made by methods known in the art. Traditional production of full-length bispecific antibodies is based on the coexpression of two immunoglobulin heavy chain-light chain pairs, where the two chains have different specificities (Milstein et al., 1983, *Nature* 305:537-539). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of 10 different antibody molecules, of which only one has the

correct bispecific structure. Purification of the correct molecule, which is usually done by affinity chromatography steps, is rather cumbersome, and the product yields are low. Similar procedures are disclosed in WO 93/08829, and in Traunecker et al., 1991, EMBO J. 10:3655-3659.

5

According to a different and more preferred approach, antibody variable domains with the desired binding specificities (antibody-antigen combining sites) are fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light chain binding, present in at least one of the fusions. DNAs encoding the immunoglobulin heavy chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. This provides for great flexibility in adjusting the mutual proportions of the three polypeptide fragments in 10 embodiments when unequal ratios of the three polypeptide chains used in the construction provide the optimum yields. It is, however, possible to insert the coding sequences for two or 15 all three polypeptide chains in one expression vector when the expression of at least two polypeptide chains in equal ratios results in high yields or when the ratios are of no particular significance.

20

In a preferred embodiment of this approach, the bispecific antibodies are composed of a hybrid immunoglobulin heavy chain with a first binding specificity in one arm, and a hybrid immunoglobulin heavy chain-light chain pair (providing a second binding specificity) in the other arm. It was found that this asymmetric structure facilitates the separation of the desired 25 bispecific compound from unwanted immunoglobulin chain combinations, as the presence of an immunoglobulin light chain in only one half of the bispecific molecule provides for a facile way of separation. This approach is disclosed in WO 94/04690. For further details for generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 1986, 121:210.

30

The invention provides functionally active fragments, derivatives or analogs of the anti-polypeptide immunoglobulin molecules. Functionally active means that the fragment, derivative or analog is able to elicit anti-anti-idiotype antibodies (i.e., tertiary antibodies) that recognize the same antigen that is recognized by the antibody from which the fragment, 35 derivative or analog is derived. Specifically, in a preferred embodiment the antigenicity of the idiotype of the immunoglobulin molecule may be enhanced by deletion of framework and CDR sequences that are C-terminal to the CDR sequence that specifically recognizes the antigen. To determine which CDR sequences bind the antigen, synthetic peptides containing the CDR sequences can be used in binding assays with the antigen by any binding assay 40 method known in the art.

The present invention provides antibody fragments such as, but not limited to,  $F(ab')^2$  fragments and Fab fragments. Antibody fragments which recognize specific epitopes may be generated by known techniques.  $F(ab')^2$  fragments consist of the variable region, the light

chain constant region and the CH1 domain of the heavy chain and are generated by pepsin digestion of the antibody molecule. Fab fragments are generated by reducing the disulfide bridges of the F(ab')<sup>2</sup> fragments. The invention also provides heavy chain and light chain dimmers of the antibodies of the invention, or any minimal fragment thereof such as Fvs or 5 single chain antibodies (SCAs) (e.g., as described in U.S. 4,946,778; Bird, 1988, Science 242:423-42; Huston et al., 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; and Ward et al., 1989, Nature 334:544-54), or any other molecule with the same specificity as the antibody of the invention. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

10 Techniques for the assembly of functional Fv fragments in *E. coli* may be used (Skerra et al., 1988, Science 242:1038-1041).

In other embodiments, the invention provides fusion proteins of the immunoglobulins of the invention (or functionally active fragments thereof), for example in which the 15 immunoglobulin is fused via a covalent bond (e.g., a peptide bond), at either the N-terminus or the C-terminus to an amino acid sequence of another protein (or portion thereof, preferably at least 10, 20 or 50 amino acid portion of the protein) that is not the immunoglobulin. Preferably the immunoglobulin, or fragment thereof, is covalently linked to the other protein at the N-terminus of the constant domain. As stated above, such fusion proteins may facilitate 20 purification, increase half-life in vivo, and enhance the delivery of an antigen across an epithelial barrier to the immune system.

The immunoglobulins of the invention include analogs and derivatives that are either modified, i.e., by the covalent attachment of any type of molecule as long as such covalent 25 attachment that does not impair immunospecific binding. For example, but not by way of limitation, the derivatives and analogs of the immunoglobulins include those that have been further modified, e.g., by glycosylation, acetylation, pegylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other protein, etc. Any of numerous chemical modifications may be carried out by 30 known techniques, including, but not limited to specific chemical cleavage, acetylation, formylation, etc. Additionally, the analog or derivative may contain one or more non-classical amino acids.

The foregoing antibodies can be used in methods known in the art relating to the localization 35 and activity of the polypeptides of the invention, e.g., for imaging or radioimaging these proteins, measuring levels thereof in appropriate physiological samples, in diagnostic methods, etc. and for radiotherapy.

The antibodies of the invention can be produced by any method known in the art for the 40 synthesis of antibodies, in particular, by chemical synthesis or by recombinant expression, and are preferably produced by recombinant expression technique.

Recombinant expression of antibodies, or fragments, derivatives or analogs thereof, requires construction of a nucleic acid that encodes the antibody. If the nucleotide sequence of the

antibody is known, a nucleic acid encoding the antibody may be assembled from chemically synthesized oligonucleotides (e.g., as described in Kutmeier et al., 1994, *BioTechniques* 17:242), which, briefly, involves the synthesis of overlapping oligonucleotides containing portions of the sequence encoding antibody, annealing and ligation of those oligonucleotides, and then amplification of the ligated oligonucleotides by PCR.

5 Alternatively, the nucleic acid encoding the antibody may be obtained by cloning the antibody. If a clone containing the nucleic acid encoding the particular antibody is not available, but the sequence of the antibody molecule is known, a nucleic acid encoding the antibody may be obtained from a suitable source (e.g., an antibody cDNA library, or cDNA library generated from any tissue or cells expressing the antibody) by PCR amplification using synthetic primers hybridizable to the 3' and 5' ends of the sequence or by cloning using an oligonucleotide probe specific for the particular gene sequence.

10 15 If an antibody molecule that specifically recognizes a particular antigen is not available (or a source for a cDNA library for cloning a nucleic acid encoding such an antibody), antibodies specific for a particular antigen may be generated by any method known in the art, for example, by immunizing an animal, such as a rabbit, to generate polyclonal antibodies or, more preferably, by generating monoclonal antibodies. Alternatively, a clone encoding at least the Fab portion of the antibody may be obtained by screening Fab expression libraries (e.g., as described in Huse et al., 1989, *Science* 246:1275-1281) for clones of Fab fragments that bind the specific antigen or by screening antibody libraries (See, e.g., Clackson et al., 20 1991, *Nature* 352:624; Hane et al., 1997 *Proc. Natl. Acad. Sci. USA* 94:4937).

25 25 Once a nucleic acid encoding at least the variable domain of the antibody molecule is obtained, it may be introduced into a vector containing the nucleotide sequence encoding the constant region of the antibody molecule (see, e.g., WO 86/05807; WO 89/01036; and U.S. 5,122,464). Vectors containing the complete light or heavy chain for co-expression with the nucleic acid to allow the expression of a complete antibody molecule are also available.

30 30 Then, the nucleic acid encoding the antibody can be used to introduce the nucleotide substitution(s) or deletion(s) necessary to substitute (or delete) the one or more variable region cysteine residues participating in an intrachain disulfide bond with an amino acid residue that does not contain a sulphydryl group. Such modifications can be carried out by any method known in the art for the introduction of specific mutations or deletions in a nucleotide sequence, for example, but not limited to, chemical mutagenesis, in vitro site directed 35 mutagenesis (Hutchinson et al., 1978, *J. Biol. Chem.* 253:6551), PCR based methods, etc.

40 In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., 1984, *Proc. Natl. Acad. Sci.* 81:851-855; Neuberger et al., 1984, *Nature* 312:604-608; Takeda et al., 1985, *Nature* 314:452-454) by splicing genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. As described supra, a chimeric antibody is a molecule in which different portions are derived from different animal species, such as those

having a variable region derived from a murine mAb and a human antibody constant region, e.g., humanized antibodies.

Once a nucleic acid encoding an antibody molecule of the invention has been obtained, the 5 vector for the production of the antibody molecule may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing the protein of the invention by expressing nucleic acid containing the antibody molecule sequences are described herein. Methods which are well known to those skilled in the art can be used to construct expression vectors containing an antibody molecule coding sequences and 10 appropriate transcriptional and translational control signals. These methods include, for example, in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. See, for example, the techniques described in Sambrook et al. (1990, Molecular Cloning, A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) and Ausubel et al. (eds., 1998, Current Protocols in Molecular Biology, 15 John Wiley & Sons, NY).

The expression vector is transferred to a host cell by conventional techniques and the transfected cells are then cultured by conventional techniques to produce an antibody of the invention.

20 The host cells used to express a recombinant antibody of the invention may be either bacterial cells such as *Escherichia coli*, or, preferably, eukaryotic cells, especially for the expression of whole recombinant antibody molecule. In particular, mammalian cells such as Chinese hamster ovary cells (CHO); in conjunction with a vector such as the major intermediate early 25 gene promoter element from human cytomegalovirus is an effective expression system for antibodies (Foecking et al., 198, Gene 45:101; Cockett et al., 1990, Bio/Technology 8:2).

30 A variety of host-expression vector systems may be utilized to express an antibody molecule of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, express the antibody molecule of the invention *in situ*. These include but are not limited to microorganisms such as bacteria (e.g., *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors 35 containing antibody coding sequences; yeast (e.g., *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing antibody coding sequences; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing the antibody coding sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed 40 with recombinant plasmid expression vectors (e.g., Ti plasmid) containing antibody coding sequences; or mammalian cell systems (e.g., COS, CHO, BHK, HEK 293, 3T3 cells) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the antibody molecule being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions comprising an antibody molecule, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther et al., 1983, EMBO J. 2:1791), in which the antibody coding sequence may be ligated individually into the vector in frame with the lac Z coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, 1985, Nucleic Acids Res. 13:3101-3109; Van Heeke & Schuster, 1989, J. Biol. Chem. 24:5503-5509); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption and binding to a matrix glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

In an insect system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The antibody coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). In mammalian host cells, a number of viral-based expression systems (e.g., an adenovirus expression system) may be utilized.

As discussed above, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein.

For long-term, high-yield production of recombinant antibodies, stable expression is preferred. For example, cells lines that stably express an antibody of interest can be produced by transfecting the cells with an expression vector comprising the nucleotide sequence of the antibody and the nucleotide sequence of a selectable (e.g., neomycin or hygromycin), and selecting for expression of the selectable marker. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that interact directly or indirectly with the antibody molecule.

The expression levels of the antibody molecule can be increased by vector amplification (for a review, see Bebbington and Hentschel, The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells in DNA cloning, Vol.3. (Academic Press, New York, 1987)). When a marker in the vector system expressing antibody is amplifiable, increase in the level of inhibitor present in culture of host cell will increase the number of copies of the marker gene. Since the amplified region is associated with the antibody gene, production of the antibody will also increase (Crouse et al., 1983, Mol. Cell. Biol. 3:257).

The host cell may be co-transfected with two expression vectors of the invention, the first vector encoding a heavy chain derived polypeptide and the second vector encoding a light chain derived polypeptide. The two vectors may contain identical selectable markers which enable equal expression of heavy and light chain polypeptides. Alternatively, a single vector may be used which encodes both heavy and light chain polypeptides. In such situations, the light chain should be placed before the heavy chain to avoid an excess of toxic free heavy chain (Proudfoot, 1986, *Nature* 322:52; Kohler, 1980, *Proc. Natl. Acad. Sci. USA* 77:2197). The coding sequences for the heavy and light chains may comprise cDNA or genomic DNA.

Once the antibody molecule of the invention has been recombinantly expressed, it may be purified by any method known in the art for purification of an antibody molecule, for example, by chromatography (e.g., ion exchange chromatography, affinity chromatography such as with protein A or specific antigen, and sizing column chromatography), centrifugation, differential solubility, or by any other standard technique for the purification of proteins.

Alternatively, any fusion protein may be readily purified by utilizing an antibody specific for the fusion protein being expressed. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht et al., 1991, *Proc. Natl. Acad. Sci. USA* 88:8972-897). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the open reading frame of the gene is translationally fused to an amino-terminal tag consisting of six histidine residues. The tag serves as a matrix binding domain for the fusion protein. Extracts from cells infected with recombinant vaccinia virus are loaded onto  $\text{Ni}^{2+}$  nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

In a preferred embodiment, antibodies of the invention or fragments thereof are conjugated to a diagnostic or therapeutic moiety. The antibodies can be used for diagnosis or to determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, radioactive nuclides, positron emitting metals (for use in positron emission tomography), and nonradioactive paramagnetic metal ions. See generally U.S. 4,741,900 for metal ions which can be conjugated to antibodies for use as diagnostics according to the present invention. Suitable enzymes include horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; suitable prosthetic groups include streptavidin, avidin and biotin; suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride and phycoerythrin; suitable luminescent materials include luminol; suitable bioluminescent materials include luciferase, luciferin, and aequorin; and suitable radioactive nuclides include  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{111}\text{In}$  and  $^{99}\text{Tc}$ .

Antibodies of the invention or fragments thereof can be conjugated to a therapeutic agent or drug moiety to modify a given biological response. The therapeutic agent or drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such 5 proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumour necrosis factor,  $\alpha$ -interferon,  $\beta$ -interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator, a thrombotic agent or an anti-angiogenic agent, e.g., angiostatin or endostatin; or, a biological response modifier such as a lymphokine, interleukin-1 (IL-1), interleukin-2 (IL-2), interleukin-6 (IL-6), 10 granulocyte macrophage colony stimulating factor (GM-CSF), granulocyte colony stimulating factor (G-CSF), nerve growth factor (NGF) or other growth factor.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, e.g., Arnon et al., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in 15 Monoclonal Antibodies And Cancer Therapy, Reisfeld et al. (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom et al., "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson et al. (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in Monoclonal Antibodies '84: 20 Biological And Clinical Applications, Pinchera et al. (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin et al. (eds.), pp. 303-16 (Academic Press 1985), and Thorpe et al., "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", Immunol. Rev., 62:119-58 (1982).

25 Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described in U.S. 4,676,980.

An antibody with or without a therapeutic moiety conjugated to it can be used as a therapeutic 30 that is administered alone or in combination with cytotoxic factor(s) and/or cytokine(s).

35 A seventh aspect of the invention provides methods of screening for agents that modulate (e.g., upregulate or downregulate) a characteristic of, e.g., the expression or the enzymatic or binding activity, of a polypeptide of the invention.

The invention provides methods for identifying agents (e.g., chemical compounds, 40 proteins, or peptides) that bind to a polypeptide of the invention or have a stimulatory or inhibitory effect on the expression or activity of a polypeptide of the invention. Examples of candidate agents, include, but are not limited to, nucleic acids (e.g., DNA and RNA), carbohydrates, lipids, proteins, peptides, peptidomimetics, agonists, antagonists, small molecules and other drugs. Agents can be obtained using any of the numerous suitable approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide

oligomer or small molecule libraries of compounds (Lam, 1997, *Anticancer Drug Des.* 12:145; U.S. 5,738,996; and U.S. 5,807,683, each of which is incorporated herein in its entirety by reference).

5 Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt et al., 1993, *Proc. Natl. Acad. Sci. USA* 90:6909; Erb et al., 1994, *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann et al., 1994, *J. Med. Chem.* 37:2678; Cho et al., 1993, *Science* 261:1303; Carrell et al., 1994, *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell et al., 1994, *Angew. Chem. Int. Ed. Engl.* 33:2061; and Gallop et al., 1994, *J. Med. Chem.* 37:1233, each of which is incorporated herein in its entirety by reference.

10 Libraries of compounds may be presented, *e.g.*, presented in solution (*e.g.*, Houghten, 1992, *Bio/Techniques* 13:412-421), or on beads (Lam, 1991, *Nature* 354:82-84), chips (Fodor, 1993, *Nature* 364:555-556), bacteria (U.S. Patent No. 5,223,409), spores (Patent Nos. 5,571,698; 5,403,484; and 5,223,409), plasmids (Cull et al., 1992, *Proc. Natl. Acad. Sci. USA* 89:1865-1869) or phage (Scott and Smith, 1990, *Science* 249:386-390; Devlin, 1990, *Science* 249:404-406; Cwirla et al., 1990, *Proc. Natl. Acad. Sci. USA* 87:6378-6382; and Felici, 1991, *J. Mol. Biol.* 222:301-310), each of which is incorporated herein in its entirety by reference.

15 In one embodiment, agents that interact with (i.e., bind to) a polypeptide of the invention are identified in a cell-based assay system. In accordance with this embodiment, cells expressing polypeptide of the invention are contacted with a candidate agent or a control agent and the ability of the candidate agent to interact with the polypeptide is determined. If desired, this assay may be used to screen a plurality (*e.g.* a library) of candidate agents. The cell, for example, can be of prokaryotic origin (*e.g.*, *E. coli*) or eukaryotic origin (*e.g.*, yeast or mammalian). Further, the cells can express the polypeptide of the invention endogenously or be genetically engineered to express the polypeptide. In some embodiments, polypeptide of the invention or the candidate agent is labelled, for example with a radioactive label (such as 20  $^{32}\text{P}$ ,  $^{35}\text{S}$  or  $^{125}\text{I}$ ) or a fluorescent label (such as fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde or fluorescamine) to enable detection of an interaction between a polypeptide and a candidate agent. The ability of the candidate agent to interact directly or indirectly with the polypeptide of the invention can be 25 determined by methods known to those of skill in the art. For example, the interaction between a candidate agent and a polypeptide can be determined by flow cytometry, a scintillation assay, immunoprecipitation or western blot analysis.

30 In another embodiment, agents that interact with (i.e., bind to) a polypeptide of the invention are identified in a cell-free assay system. In accordance with this embodiment, a native or recombinant polypeptide of the invention or fragment thereof is contacted with a candidate agent or a control agent and the ability of the candidate agent to interact with the polypeptide is determined. If desired, this assay may be used to screen a plurality (*e.g.* a library) of candidate agents. Preferably, the polypeptide is first immobilized, by, for example, 35 contacting the polypeptide with an immobilized antibody which specifically recognizes and binds it, or by contacting a purified preparation of polypeptide with a surface designed to bind proteins. The polypeptide may be partially or completely purified (*e.g.*, partially or 40 completely free of other polypeptides) or part of a cell lysate. Further, the polypeptide may be a fusion protein comprising the polypeptide of the invention or a biologically active portion thereof and a domain such as glutathione-S-transferase. Alternatively, the

polypeptide can be biotinylated using techniques well known to those of skill in the art (e.g., biotinylation kit, Pierce Chemicals; Rockford, IL). The ability of the candidate agent to interact with the polypeptide can be determined by methods known to those of skill in the art.

5 In another embodiment, a cell-based assay system is used to identify agents that bind to or modulate the activity of a protein, such as an enzyme, or a biologically active portion thereof, which is responsible for the production or degradation of the polypeptide of the invention or is responsible for the post-translational modification of the polypeptide. In a primary screen, a plurality (e.g., a library) of agents are contacted with cells that naturally or 10 recombinantly express: (i) a polypeptide of the invention; and (ii) a protein that is responsible for processing of the polypeptide in order to identify compounds that modulate the production, degradation, or post-translational modification of the polypeptide. If desired, agents identified in the primary screen can then be assayed in a secondary screen against cells naturally or recombinantly expressing the specific polypeptide of interest. The ability of the 15 candidate agent to modulate the production, degradation or post-translational modification of a polypeptide can be determined by methods known to those of skill in the art, including without limitation, flow cytometry, a scintillation assay, immunoprecipitation and western blot analysis.

20 In another embodiment, agents that competitively interact with (i.e., bind to) a polypeptide of the invention are identified in a competitive binding assay. In accordance with this embodiment, cells expressing the polypeptide are contacted with a candidate agent and an agent known to interact with the polypeptide; the ability of the candidate agent to competitively interact with the polypeptide is then determined. Alternatively, agents that 25 competitively interact with (i.e., bind to) a polypeptide are identified in a cell-free assay system by contacting the polypeptide with a candidate agent and an agent known to interact with the polypeptide. As stated above, the ability of the candidate agent to interact with a polypeptide of the invention can be determined by methods known to those of skill in the art. These assays, whether cell-based or cell-free, can be used to screen a plurality (e.g., a library) of candidate agents.

30 In another embodiment, agents that modulate (i.e., upregulate or downregulate) the expression of a polypeptide of the invention are identified by contacting cells (e.g., cells of prokaryotic origin or eukaryotic origin) expressing the polypeptide with a candidate agent or a control agent (e.g., phosphate buffered saline (PBS)) and determining the expression of the polypeptide or mRNA encoding the polypeptide. The level of expression of a selected 35 polypeptide or mRNA encoding polypeptide in the presence of the candidate agent is compared to the level of expression of the polypeptide or mRNA encoding the polypeptide in the absence of the candidate agent (e.g., in the presence of a control agent). The candidate agent can then be identified as a modulator of the expression of the polypeptide based on this comparison. For example, when expression of the polypeptide or mRNA encoding the 40 polypeptide is significantly greater in the presence of the candidate agent than in its absence, the candidate agent is identified as a stimulator of expression of the polypeptide or mRNA encoding the polypeptide. Alternatively, when expression of the polypeptide or mRNA encoding the polypeptide is significantly less in the presence of the candidate agent than in its absence, the candidate agent is identified as an inhibitor of the expression of the polypeptide.

or mRNA encoding the polypeptide. The level of expression of a polypeptide of the invention or the mRNA that encodes it can be determined by methods known to those of skill in the art based on the present description. For example, mRNA expression can be assessed by Northern blot analysis or RT-PCR, and protein levels can be assessed by western blot analysis.

5 In another embodiment, agents that modulate the activity of a polypeptide are identified by contacting a preparation containing the polypeptide, or cells (e.g., prokaryotic or eukaryotic cells) expressing the polypeptide with a candidate agent or a control agent and determining the ability of the candidate agent to modulate (e.g., stimulate or inhibit) the activity of polypeptide. The activity of a polypeptide can be assessed by detecting its effect on a "downstream effector" for example, but without limitation, induction of a cellular signal transduction pathway of the polypeptide (e.g., intracellular  $\text{Ca}^{2+}$ , diacylglycerol, IP3, etc.), detecting catalytic or enzymatic activity of the target on a suitable substrate, detecting the induction of a reporter gene (e.g., a regulatory element that is responsive to a polypeptide of the invention and is operably linked to a nucleic acid encoding a detectable marker, e.g., luciferase), or detecting a cellular response, for example, cellular differentiation, or cell proliferation as the case may be, based on the present description, techniques known to those of skill in the art can be used for measuring these activities (see, e.g., U.S. 5,401,639, which is incorporated in its entirety herein by reference). The candidate agent can then be identified as 10 a modulator of the activity of a polypeptide of the invention by comparing the effects of the candidate agent to the control agent. Suitable control agents include phosphate buffered 15 saline (PBS) and normal saline (NS).

20 In another embodiment, agents that modulate (i.e., upregulate or downregulate) the expression, activity or both the expression and activity of a polypeptide of the invention are 25 identified in an animal model. Examples of suitable animals include, but are not limited to, mice, rats, rabbits, monkeys, guinea pigs, dogs and cats. Preferably, the animal used represents a model of breast or kidney cancer. In accordance with this embodiment, the candidate agent or a control agent is administered (e.g., orally, rectally or parenterally such as intraperitoneally or intravenously) to a suitable animal and the effect on the expression, 30 activity or both expression and activity of the polypeptide is determined. Changes in the expression of a polypeptide can be assessed by any suitable method described above, based on the present description.

35 In yet another embodiment, a polypeptide or the invention is used as a "bait protein" in a two-hybrid assay or three hybrid assay to identify other proteins that bind to or interact with the polypeptide (see, e.g., U.S. 5,283,317; Zervos et al. (1993) Cell 72:223-232; Madura et al. (1993) J. Biol. Chem. 268:12046-12054; Bartel et al. (1993) Bio/Techniques 14:920-924; Iwabuchi et al. (1993) Oncogene 8:1693-1696; and WO 94/10300). As those skilled in the art will appreciate, such binding proteins are also likely to be involved in the propagation of signals by the polypeptides of the invention as, for example, upstream or 40 downstream elements of a signalling pathway involving the polypeptides of the invention.

This invention further provides novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

As used herein "active agent" refers to the polypeptides of the invention and nucleic acid molecules encoding the polypeptides, antibodies against the polypeptides and agents which modulate the expression of the polypeptides of the invention.

5 As discussed herein, the polypeptides, nucleic acid molecules and antibodies of the invention find use in the treatment or prophylaxis of breast and/or kidney cancer. Thus, in an eight aspect, the present invention provides a pharmaceutical formulation comprising at least one polypeptide or fragment thereof, nucleic acid molecule or antibody of the invention, optionally together with one or more pharmaceutically acceptable excipients, carriers or diluents. Preferably, the pharmaceutical formulation is for use as a vaccine and so any additional components will be acceptable for vaccine use. In addition, the skilled person will appreciate that one or more suitable adjuvants may be added to such vaccine preparations.

10

15 The medicament will usually be supplied as part of a sterile, pharmaceutical composition which will normally include a pharmaceutically acceptable carrier. This pharmaceutical composition may be in any suitable form, (depending upon the desired method of administering it to a patient).

20 It may be provided in unit dosage form, will generally be provided in a sealed container and may be provided as part of a kit. Such a kit would normally (although not necessarily) include instructions for use. It may include a plurality of said unit dosage forms.

25 The pharmaceutical composition may be adapted for administration by any appropriate route, for example by the oral (including buccal or sublingual), rectal, nasal, topical (including buccal, sublingual or transdermal), vaginal or parenteral (including subcutaneous, intramuscular, intravenous or intradermal) route. Such compositions may be prepared by any method known in the art of pharmacy, for example by admixing the active ingredient with the carrier(s) or excipient(s) under sterile conditions.

30 Pharmaceutical compositions adapted for oral administration may be presented as discrete units such as capsules or tablets; as powders or granules; as solutions, syrups or suspensions (in aqueous or non-aqueous liquids; or as edible foams or whips; or as emulsions).

35 Suitable excipients for tablets or hard gelatine capsules include lactose, maize starch or derivatives thereof, stearic acid or salts thereof.

Suitable excipients for use with soft gelatine capsules include for example vegetable oils, waxes, fats, semi-solid, or liquid polyols etc.

40 For the preparation of solutions and syrups, excipients which may be used include for example water, polyols and sugars. For the preparation of suspensions, oils (e.g. vegetable oils) may be used to provide oil-in-water or water in oil suspensions.

Pharmaceutical compositions adapted for transdermal administration may be presented as discrete patches intended to remain in intimate contact with the epidermis of the recipient for a prolonged period of time. For example, the active ingredient may be delivered from the patch by iontophoresis as generally described in *Pharmaceutical Research*, 3(6):318 (1986).

5

Pharmaceutical compositions adapted for topical administration may be formulated as ointments, creams, suspensions, lotions, powders, solutions, pastes, gels, sprays, aerosols or oils. For infections of the eye or other external tissues, for example mouth and skin, the compositions are preferably applied as a topical ointment or cream. When formulated in an ointment, the active

10

ingredient may be employed with either a paraffinic or a water-miscible ointment base.

Alternatively, the active ingredient may be formulated in a cream with an oil-in-water cream base or a water-in-oil base. Pharmaceutical compositions adapted for topical administration to the eye include eye drops wherein the active ingredient is dissolved or suspended in a suitable carrier, especially an aqueous solvent. Pharmaceutical compositions adapted for topical administration

15

in the mouth include lozenges, pastilles and mouth washes.

Pharmaceutical compositions adapted for rectal administration may be presented as suppositories or enemas.

20

Pharmaceutical compositions adapted for nasal administration wherein the carrier is a solid include a coarse powder having a particle size for example in the range 20 to 500 microns which is administered in the manner in which snuff is taken, i.e. by rapid inhalation through the nasal passage from a container of the powder held close up to the nose. Suitable compositions wherein the carrier is a liquid, for administration as a nasal spray or as nasal drops, include aqueous or oil solutions of the active ingredient.

25

Pharmaceutical compositions adapted for administration by inhalation include fine particle dusts or mists which may be generated by means of various types of metered dose pressurised aerosols, nebulisers or insufflators.

30

Pharmaceutical compositions adapted for vaginal administration may be presented as pessaries, tampons, creams, gels, pastes, foams or spray formulations.

35

Pharmaceutical compositions adapted for parenteral administration include aqueous and non-

aqueous sterile injection solution which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation substantially isotonic with the blood of the intended recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents and thickening agents. Excipients which may be used for injectable solutions include water, alcohols, polyols, glycerine and vegetable oils, for example. The compositions may be

40

presented in unit-dose or multi-dose containers, for example sealed ampoules and vials, and may be stored in a freeze-dried (lyophilised) condition requiring only the addition of the sterile liquid carried, for example water for injections, immediately prior to use. Extemporaneous injection solutions and suspensions may be prepared from sterile powders, granules and tablets.

The pharmaceutical compositions may contain preserving agents, solubilising agents, stabilising agents, wetting agents, emulsifiers, sweeteners, colourants, odourants, salts (substances of the present invention may themselves be provided in the form of a pharmaceutically acceptable salt), buffers, coating agents or antioxidants. They may also contain therapeutically active agents in

5 addition to the substance of the present invention.

Dosages of the substance of the present invention can vary between wide limits, depending upon the disease or disorder to be treated, the age and condition of the individual to be treated, etc. and a physician will ultimately determine appropriate dosages to be used. This dosage may be

10 repeated as often as appropriate. If side effects develop the amount and/or frequency of the dosage can be reduced, in accordance with normal clinical practice.

In a ninth aspect, the present invention provides a method for the prophylaxis and/or treatment of breast and/or kidney cancer in a subject, which comprises administering to said 15 subject a therapeutically effective amount of at least one polypeptide or fragment thereof, nucleic acid molecule or antibody of the invention.

In a tenth aspect, the present invention provides the use of at least one polypeptide or fragment thereof, nucleic acid molecule or antibody of the invention in the preparation of a 20 medicament for use in the prophylaxis and/or treatment of breast and/or kidney cancer. In particular, the preparation of vaccines and/or compositions comprising or consisting of antibodies is a preferred embodiment of this aspect of the invention.

In view of the importance of BCMP 101 in breast and kidney cancer the following form 25 additional aspects of the present invention:

- i) a method for monitoring/assessing breast and/or kidney cancer treatment in a patient, which comprises the step of determining the presence or absence and/or quantifying at least one polypeptide, at least one nucleic acid molecule or at least one antibody of the invention in a biological sample obtained from said patient;
- ii) a method for the identification of metastatic breast and/or kidney cancer cells in a biological sample obtained from a subject, which comprises the step of determining the presence or absence and/or quantifying at least one polypeptide, at least one nucleic acid molecule or at least one antibody of the invention.

In the context of the present invention, the biological sample can be obtained from any source, such as a serum sample or a tissue sample, e.g. breast or kidney tissue. When looking 40 for evidence of metastasis, one would look at major sites of breast metastasis such as lymph nodes, liver, lung and/or bone and of kidney metastasis, such as lymph nodes, lung and/or bone.

Preferred features of each aspect of the invention are as for each of the other aspects *mutatis mutandis*. The prior art documents mentioned herein are incorporated to the fullest extent permitted by law.

5 The invention will now be described with reference to the following examples, which should not in any way be construed as limiting the scope of the present invention. The examples refer to the figures in which:

10 **Figure 1:** shows the nucleotide and predicted amino acid sequences of BCMP 101. The tandem mass spectrum is in bold and italicized. MALDI mass spectra are in bold and underlined. Note that six nucleic acid differences, which may be due to polymorphisms, resulting in 5 amino acids changes, have been identified between the cloned sequence and the predicted sequence based on the assembly of public domain ESTs and of genomic data (see Example 1). These differences are shaded, with the 15 assembled sequence nucleic/amino acid listed below the cloned sequence, which is italicized;

20 **Figure 2:** shows tissue distribution of BCMP 101 mRNA. Levels of mRNA in normal tissues (including kidney) and two kidney cancer cell lines (Wilm's tumour cell line G-401 and human embryonic kidney cell line 293) were quantified by real time RT-PCR. mRNA levels are expressed as the number of copies  $\text{ng}^{-1}$  cDNA. Note the 40-fold difference between the vertical scale used for kidney cancer cell lines and the vertical scale used for normal tissues;

25 **Figure 3:** shows the expression of BCMP 101 in normal and tumour breast tissues. Levels of BCMP 101 mRNA in matched normal and tumour tissues from seven breast cancer patients were measured by real time RT-PCR. mRNA levels are expressed as the number of copies  $\text{ng}^{-1}$  cDNA;

30 **Figure 4:** shows the expression of BCMP 101 in normal and tumour breast tissues. Samples 1-23 are tumour samples not involving metastasis to the lymph nodes. Samples 26-50 are tumour samples involving metastases to variable numbers of lymph nodes. The final 8 samples are from normal breast tissue (reduction mammoplasties). mRNA levels are expressed as the number of copies  $\text{ng}^{-1}$  cDNA. 35 There is a statistically significant difference between all tumour samples and normal samples (T-test,  $p < 0.05$ ); and

40 **Figures 5:** *in situ* RT PCR analysis of BCMP 101 expression in sections of invasive ductal breast cancer tissue (upper panel), and consecutive negative control section in which the BCMP 101 primers have been replaced with primers to a control gene (Prostate Specific Antigen) (lower panel). Note the high BCMP 101 expression (dark staining) in a portion of epithelial hyperplasia (typical of breast carcinoma), flanked with two arrowheads in the upper panel.

**Figures 6:** cellular localisation of BCMP 101 in breast cancer cell lines. Fluorescence microscopy showing expression of GFP-tagged BCMP101 in MDA-MB-468 (A) and T-47D (B) cell lines. Membrane localisation is indicated by white arrowheads. Magnification using x 60 oil immersion objective.

5

**Figures 7:** expression of BCMP 101 in breast cancer tissue sections by immunohistochemistry. Expression of BCMP 101 in carcinoma cells is indicated by arrowheads.

10 **Example 1: Identification and cloning of BCMP 101**

Protein BCMP 101 was isolated from MDA-MB-468 cell membranes.

The breast carcinoma cell line MDA-MB-468 (ATCC:HTB-132) was cultured and integral membranes were extracted with the Tx114 detergent. These were subsequently analyzed by two-dimensional gel electrophoresis as described in U.S. Patent No 6,064,754.

15

Mass Spectrometry

Proteins excised from the 2D gel were digested with trypsin and analyzed by MALDI-TOF-MS (Voyager STR, Applied Biosystems) using a 337 nm wavelength laser for desorption and the reflectron mode of analysis. Selected masses for BCMP 101 were further characterized by tandem mass spectrometry using a QTOF-MS equipped with a nanospray ion source, (Micromass UK Ltd.). Prior to MALDI analysis the samples were desalted and concentrated using C18 Zip Tips™ (Millipore). Samples for tandem MS were purified using a nano LC system (LC Packings) incorporating C18 SPE material.

Using the SEQUEST search program (Eng et al., 1994, J. Am. Soc. Mass Spectrom. 5:976-989), uninterpreted tandem mass spectra of tryptic digest peptides were searched against a database of public domain proteins constructed of protein entries in the non-redundant database held by the National Centre for Biotechnology Information (NCBI) which is accessible at <http://www.ncbi.nlm.nih.gov/> and also constructed of Expressed Sequence Tags entries (<http://www.ncbi.nlm.nih.gov/dbEST/index.html>). As a result of database searching, the following amino acid sequence of a tryptic digest peptide of BCMP 101 was determined from a match to a tryptic digest peptide in a conceptual translation of EST AI472043: NSESFAAWCR (shown in Figure 1).

EST AI472043 corresponds to base pairs 558-1054 of the DNA sequence in Figure 1. ESTs AI684699 (corresponding to bp 80-565) and AI827549 (corresponding to bp 45-547) were used to establish the full length ORF (note that AI827549 includes the in frame stop codon just upstream of the ATG). The sense primer used to amplify the full length clone was designed to genomic sequences in entry AC021396 (see below Example 3 on chromosomal localization). The identified protein had a pI of 5.3 and a MW of 39940 as measured by 2-D gel analysis as described above, the predicted pI and MW for this protein is 5.34 and 34474 respectively.

A full length clone was amplified by PCR from MDA-MB-468 cDNA:

Preparation of total RNA and cDNA synthesis

Total RNA was prepared from cultured cells and tissue samples using Trizol reagent (Life Technologies), according to the manufacturer's instructions, and resuspended in RNase-free water at a concentration of 1 $\mu$ g/ $\mu$ l. 1 to 5 $\mu$ g total RNA were used as a template for cDNA synthesis using an oligo dT primer and the Superscript II reverse transcription kit (Life Technologies). cDNAs were column purified (Qiagen) and eluted at a concentration of 10ng/ $\mu$ l.

Cloning of BCMP 101 cDNA

10 The predicted full length BCMP 101 ORF was amplified by PCR from MDA-MB-468 cDNAs, using the following primers:  
 Sense , 5' TGTGCAAATGACCCTGGAGTTG 3' (SEQ ID NO: 3);  
 Antisense, 5' GGCTGCTACTGCAAACAGTTCC 3' (SEQ ID NO: 4).

15 Reactions contained 10ng cDNA and reagents for PCR (Clontech, Advantage-GC 2 PCR kit), and used the following cycling parameters: 1 cycle of 94°C for 3 minutes, followed by 40 cycles of 94°C for 30s, 65°C for 30 seconds, 72°C for 90 seconds. The PCR products were column purified (Qiagen), cloned into a T/A vector (Invitrogen) and the nucleotide sequence subsequently verified (University of Oxford, Sequencing Facility, UK).

20 The BCMP 101 sequence (Figure 1, SEQ ID NO: 1) matches the following GenBank entry (available at: <http://www.ncbi.nlm.nih.gov/>): CAD10629 - NSE2 protein [Homo sapiens] - A novel NS-containing protein.

25 Example 2: Expression of BCMP 101 mRNA in human tissues

We used real time quantitative RT-PCR (Heid, C.A., Stevens, J., Livak, K.J. & Williams, P.M. Real time quantitative PCR. *Genome Res.* **6**, 986-994 (1996); Morrison, T.B., Weis, J.J. & Wittwer, C.T. Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *Biotechniques* **24**, 954-958 (1998)) to analyse the distribution of BCMP 101 mRNA in normal human tissues and kidney cancer cell lines (Fig 2). Note the 40-fold difference between the right-hand scale, used for kidney cancer cell lines, and the left-hand scale, used for normal human tissues, which includes normal kidney.

35 Quantification of BCMP 101 mRNA by RT-PCR

Real time RT-PCR was used to quantitatively measure BCMP 101 expression in normal human tissue mRNAs (Clontech), kidney cancer cell line mRNAs (Ambion), breast cancer tissues removed during surgery, and normal breast tissue removed during breast reduction mammoplasty. Ethical approval for the normal and tumour breast samples was obtained at surgery (University of Oxford, UK). The primers used for PCR were as follows:

sense, 5' GGTCAACGATCTGTACCGCTAC 3' (SEQ ID NO: 5),  
 antisense, 5' GCCGATCTTGAACTCGCGCTTG 3' (SEQ ID NO: 6).

Reactions containing 10ng cDNA, prepared as described above, SYBR green sequence detection reagents (PE Biosystems) and sense and antisense primers were assayed on an ABI7700 sequence detection system (PE Biosystems). The PCR conditions were 1 cycle at 50°C for 2 min, 1 cycle at 95°C for 10 min, and 40 cycles of 95°C for 15s, 65°C for 1min.

5 The accumulation of PCR product was measured in real time as the increase in SYBR green fluorescence, and the data were analyzed using the Sequence Detector program v1.6.3 (PE Biosystems). Standard curves relating initial template copy number to fluorescence and amplification cycle were generated using the amplified PCR product as a template, and were used to calculate BCMP 101 copy number in each sample.

10 The distribution of BCMP 101 mRNA was low in normal tissues, with the highest levels of expression in mammary gland, kidney and bladder (130-240 copies ng<sup>-1</sup> cDNA). BCMP 101 mRNA was detected at a high level in two kidney cancer cell lines, Human Embryonic Kidney cell line 293 and Wilm's tumour G-401 cell line (3300 and 11,000 copies ng<sup>-1</sup> cDNA respectively).

15 The distribution of BCMP 101 mRNA in clinical breast carcinomas was measured in matched normal and tumour tissue samples from seven breast cancer patients (Fig 3). BCMP 101 expression was increased in all the tumour samples, relative to their matched normal tissues, 20 with four of the samples showing a 4- fold or more elevation in expression. Furthermore, the difference between the paired normal and tumour sample sets was highly statistically significant, with a p-value of 0.014. Thus, BCMP 101 shows a restricted pattern of expression in normal human tissues, and is elevated in some breast tumours, suggesting that this protein has potential as a therapeutic target.

25 To further examine the expression of this gene in breast cancer tissues, we extended the quantification of BCMP 101 mRNA levels to a further set of 40 tumour samples (Figure 4). BCMP 101 expression level was found here again to be significantly correlated with the presence of breast tumour.

30 **Example 3: Chromosomal localisation**  
 A Blast search (<http://www.ncbi.nlm.nih.gov/BLAST/>) with the BCMP 101 cDNA sequence (Figure 1) against htgs (High-Throughput Genome Sequences), returns the following GenBank clone: AC021396, mapped to chromosome 8q23.  
 35 Furthermore, a gene responsible for Polycystic Kidney Disease (PKD) in a rat model of autosomal dominant PKD has been identified on rat chromosome 5 (Location of the first genetic locus, PKDr1, controlling autosomal dominant polycystic kidney disease in Han:SPRD cy/+ rat, Bihoreau MT, Ceccherini I, Browne J, Kranzlin B, Romeo G, Lathrop GM, James MR, Gretz N., Hum Mol Genet 1997 Apr;6(4):609-13). A detailed linkage 40 mapping of rat chromosome 5 placed this PKD locus about 25 cM from the proenkephalin gene, which in human is located on 8q23-q24.

**Example 4: In situ RT-PCR**

To further illustrate the involvement of BCMP101 in breast cancer, in situ RT-PCR analysis of BCMP101 expression has been carried out on sections of invasive ductal breast cancers.

5 Formalin fixed, paraffin embedded breast tissue from a patient with ductal carcinoma was sectioned (5 microM thick) onto glass slides (provided by Human Research Tissue Bank, Department of Cellular Pathology, Peterborough District Hospital, Thorpe Road, Peterborough PE3 6DA). Briefly, the tissue was de-waxed in xylene, gradually re-hydrated through alcohol and washed in phosphate buffered saline (PBS) before being permeabilised in  
 10 0.01% Triton X-100 for 3 minutes followed by treatment with Proteinase K for 30 minutes at 37°C.

Direct *in situ* RT PCR was carried out in a GeneAmp In Situ PCR System 1000 (Perkin Elmer Biosystems) using a GeneAmp Thermostable rTth RT PCR kit (Perkin Elmer Biosystems).

15 The primers used to amplify BCMP101 were as follows:

sense, 5'-TTCACCTCTCCGCGGGTAGCCT-3' (SEQ ID NO: 7),  
 antisense , 5'-GGAAGTTACCCACATATACGGC-3' (SEQ ID NO: 8).

20 The thermal cycling parameters were; 1 cycle of 94°C for 2.5 minutes followed by 20 cycles of 94°C for 40 seconds, 60°C for 50 seconds, and 72°C for 30 seconds. Amplified product was detectable through the direct incorporation of alkali stable Digoxigenin-11-dUTP (Roche Diagnostics Ltd.) which was added to the reaction mix. After washing in PBS an anti-Digoxigenin-Gold antibody (Roche Diagnostics Ltd.) was incubated on the tissue section for  
 25 30 minutes at room temperature, this was followed by a silver enhancement step (Roche Diagnostics Ltd. silver enhancement reagents) during which time the amplified expression product became visible by light microscopy. The tissue was counter-stained with hematoxylin (Dako Ltd.) and images were captured by a digital camera attached to a light microscope (x10 objective).

30 In the paired images of Figure 5 the upper panel demonstrates BCMP 101 expression in the breast cancer tissue, the lower panel represents a negative control consecutive section where the BCMP101 primers have been replaced with primers to a control gene (Prostate Specific Antigen). It is clearly apparent from these Figures that BCMP 101 is specifically expressed in  
 35 the cancerous ductal epithelial cells of this breast cancer tissue (compare with surrounding breast tissue and negative control experiment). For example, a portion of epithelial hyperplasia (typical of breast carcinoma) has been flanked with two arrowheads (upper panel); this shows that the area of dark staining (representing BCMP101 expression) is restricted to the cancer cells.

40

**Example 5: Cellular Localisation of BCMP101 in Breast Cancer Cell Lines**

C-terminal tagging with SuperGloTM green fluorescent protein (GFP) was used to determine the cellular localisation of BCMP101 in MDA-MB-468 and T-47D cell lines.

The BCMP101 full length ORF was PCR cloned into the pQBI25/50-fN1 vector (Qbiogene) resulting in an in-frame addition of the SuperGloTM(sg)GFP protein to the C-terminus of each expressed protein. Transient transfection of sgGFP-tagged BCMP101 cDNA into MDA-MB-468 and T-47D cell lines was achieved using SuperfectTM transfection reagent (Qiagen) according to the manufacturers instructions. Transfected cells were washed in phosphate buffered saline (PBS), fixed in 4% paraformaldehyde for 30 minutes, then washed again in PBS before being mounted in an aqueous-based fluorescent mounting medium (Dako Ltd.). Fluorescence images were captured using a DC300F digital camera attached to a DMIRE2 fluorescence microscope (Leica Microsystems (UK) Ltd.)

Analysis of the cellular location of GFP-tagged BCMP101 demonstrated high expression in the Golgi and endoplasmic reticulum as well as significant expression associated with the plasma membrane in both MDA-MB-468 and T-47D cell lines (Fig. 6). Indeed, we observed particularly high levels of BCMP101 plasma membrane localisation in areas of cell-cell contact (Fig. 6B)). These data suggest a role for BCMP101 in the transduction of a cellular signal mediated by cell-cell contact

#### **Example 6: BCMP101 Expression in Breast Cancer Tissues by Immunohistochemistry**

Immunohistochemical analysis was used to determine the expression of BCMP101 in sections of breast cancer tissues. Immunohistochemical analysis was carried out on frozen section of a breast tumour (from Peterborough Tissue Bank, ref. No 6574 - Human Research Tissue Bank, Department of Cellular Pathology, Peterborough District Hospital, Thorpe Road, Peterborough PE3 6DA - <http://www.tissuebank.co.uk/>). Frozen sections were thawed at room temperature for 30 min, and fixed in acetone at 4 degrees C for 10 min. The sections were subsequently washed twice in PBS. Endogenous hydrogen peroxidase activity was quenched by treating the slides in 3% hydrogen peroxidase/PBS for 10 mins, followed by 2 washes in PBS. The tissue was blocked in 10% donkey serum/PBS for 1 hour before addition of 2 microg/ml primary polyclonal antibody (in 2.5% donkey serum). The BCMP101 polyclonal antibody was raised in rabbits immunised with a specific peptide (Abcam Ltd., Cambridge,UK). The peptide sequence was chosen for synthesis based on plots of hydrophobicity, antigenicity, surface probability. The peptide was synthesised using Fmoc chemistry with a cysteine residue added to the end to enable specific thiol reactive coupling of Keyhole Limpet Haemocyanin prior to immunisation. The BCMP101 peptide used was: SYKEVPTADPTGVDR. Western blot analysis of T-47D and MDA-MB-468 cell lysates was used to confirm that the antibody cross-reacted with a single band of the predicted size. Following 3 washes in PBS the tissue sections were incubated with biotin conjugated secondary antibodies (Biotin-SP-conjugated AffiniPure™ Donkey anti-rabbit, Jackson ImmunoResearch) diluted at 1:200 (2.5 microg/ml in 2.5% donkey serum/PBS) for 1 hour. Slides were washed 3 times in PBS and the tissue incubated with Streptavidin-HRP (Jackson ImmunoResearch) diluted 1:100 (5

microg/ml in 2.5% donkey serum/PBS), followed by three 5 min washes in PBS. Antibody signal was detected using DAB substrate solution (Dako Ltd.) according to the manufacturers instructions.

5 Immunohistochemical analysis of BCMP101 expression demonstrated very low levels in multiple normal tissues. In contrast, high levels of BCMP101 (Fig. 7) were detected in the carcinoma cells of multiple breast cancer tissues.

**Example 7: Yeast Two-Hybrid Study of BCMP 101 Interacting factors**

10 Most biological processes involve specific protein-protein interactions. Protein-protein interactions which may be transient or otherwise, enable two or more proteins or subunits to associate. Protein-protein interactions can have a number of measurable effects: they can alter kinetic properties of proteins, they are a common mechanism for allowing substrate channeling; they can result in the formation of new binding sites, they can alter the activity of 15 a protein and protein-protein interactions can alter the specificity of a protein for its substrate (Phizicky and Fields, 1995, Microbiological Reviews, 59: 94-123; Pawson and Nash, 2000, Gene Dev. 14: 1027-1047).

20 General methodologies to identify interacting proteins or to study these interactions have been developed. Among these methods is the two-hybrid system originally developed by Fields and co-workers and described, for example, in U.S. 5,283,173, U.S. 5,468,614 and U.S. 5,667,973. A summary of the available methodologies for detecting protein-protein interactions is described in Vidal and Legrain, *Nucleic Acids Research* Vol. 27, No. 4 pgs. 919-929 (1999) and Legrain and Selig, *FEBS Letters* 480 pgs. 32-36 (2000), which references 25 are incorporated herein by reference.

30 One method to overcome the difficulties encountered with the methods in the prior art is described in WO 99/42612, incorporated herein by reference. This method is similar to the two-hybrid system described in the prior art in that it also uses bait and prey polypeptides.

35 The method described in WO99/42612 permits the screening of more prey polynucleotides with a given bait polynucleotide in a single step than in the prior art systems due to the cell to cell mating strategy between haploid yeast cells. Furthermore, this method is more thorough and reproducible, as well as sensitive. Thus, the presence of false negatives and/or false positives is extremely minimal as compared to the conventional prior art methods.

*Construction of prey plasmids*

40 Random-primed cDNA was prepared, ligated into an appropriate vector using methods known in the art, and transformed into electrocompetent cells as described in WO 00/66722 and WO 02/12290.

cDNA was collected from the cells (above) and transformed into yeast strain YHGX13 (MATalpha Gal4delta Gal80delta ade2-101::KAN<sup>R</sup>, his3, leu2-3, -112, trp1-901, ura3-52 URA3::UASGAL1-LacZ, Met) as described in the above Patent Applications.

*Construction of bait plasmids*

Bait fragments were cloned into appropriate vectors using methods known in the art.

5 Amplification, PCR and screening the collection with the two-hybrid system in yeast was as described in WO 00/66722. Where the number of His+ cell clones <285 the process stamp overlay protocol below is used on all colonies. If the number of His+ cell clones >285 and <5000: then process via the overlay and then the stamp overlay protocols on all blue colonies (see WO 0066722 sections 2.B and 2.C).

10

*The stamp overlay assay*

His+ colonies were grown overnight at 30°C in microtiter plates containing DO-Leu-Trp-His-Tetracyclin medium with shaking. The day after the overnight culture, the 96 colonies were stamped on a 15cm plate of DO-Leu-Trp-His. 4 control yeast colonies were spotted on the same plate. After 2 days of growing at 30°C, an overlay assay was performed on this plate with 80ml of overlay mixture (see WO 00/66722 section 2.B.). After 2 hours of incubation, the plate was photographed with a CCD camera. The blue intensity was quantified by Genetools™ software (SYNGENE) and normalized to the control spots.

15 Positive clones were identified and plasmids rescued and transformed into electrocompetent cells as described in WO 02/12290 and WO 00/66722.

20 *Protein-protein interactions*  
Identification of prey nucleotide sequences was as described (WO 02/12290 and WO 00/66722) or alternatively, prey nucleotide sequences can be compared with one another and 25 those which share identity over a significant region (60nt) can be grouped together to form a contiguous sequence (Contig) whose identity can be ascertained in the same manner as for individual prey fragments (WO 02/12290 and WO 00/66722). Selected Interacting Domains can be determined also as described in the latter applications.

30 *Major BCMP101 Interacting Factors*

BCMP101 interacts with alpha-1 catenin (Swiss-Prot accession P35221) and membrane components of clathrin coated vesicles AP1M2 (GenBank accession NP\_005489) and AP47 (GenBank accession NP\_115882).

35 Alpha-1 catenin associates with the cytoplasmic domains of multiple plasma membrane localised cadherins and as such are thought to play an important role in cell-cell adhesion. Interestingly, alpha-1 catenin is mutated in the invasive human colon cancer cell family HCT-8 and is therefore an invasion suppressor gene in human colon cancer (Vermeulen, S.J., Nollet, F., Teugels, E., Vennekens, K.M., Malfait, F., Philippe, J., Speleman, F., Bracke, M.E., van Roy, F.M. & Mareel, M.M. The alphaE-catenin gene (CTNNA1) acts as an 40 invasion-suppressor gene in human colon cancer cells. *Oncogene* 18, 905-915 (1999)).

Additional evidence supports an interaction between BCMP101 and alpha-1 catenin in breast cancer cells. Firstly, alpha-1 catenin was identified in a proteomic analysis of the same breast cancer cell line (MDA-MB-468) as the one described above in Example 1. Secondly, association of BCMP101 with alpha-1 catenin accounts for the plasma membrane localisation

of GFP-tagged BCMP101 seen in breast cancer cell lines (Fig. 6A), and in particular the observation that BCMP101 was plasma membrane associated at sites of cell-cell contact (Fig. 6B). Thus BCMP101 is associated with a complex of membrane proteins with a known role in tumour progression.

5 Figure 6 also shows some cytosolic and Golgi localisation of GFP tagged BCMP101. This is consistent with the interaction of BCMP101 with AP1M2 and AP47 which are part of the AP-1 adaptor complex which is recruited from the cytosol onto the trans-Golgi network membrane, where it co-assembles with clathrin into a coat that drives vesicle budding involved in trafficking proteins to other cell membranes including the plasma membrane

## CLAIMS

1. A polypeptide which:

5 a) comprises or consists of the amino acid sequence shown in figure 1 (SEQ ID NO: 1);  
b) is a derivative having one or more amino acid substitutions, deletions or insertions relative to the amino acid sequence shown in figure 1 (SEQ ID NO: 1); or  
c) is a fragment of a polypeptide as defined in a) or b) above, which is at least ten amino acids long.

10

2 A polypeptide as claimed in claim 1, wherein the pI is about 5.3 and the MW is about 39940Da.

15 3 A polypeptide as claimed in claim 1, wherein the pI is within 10% of 5.3 and the MW is within 10% of 39940Da.

4 A polypeptide as claimed in claim 1, wherein the pI is within 5% of 5.3 and the MW is within 5% of 39940Da.

20 5. A polypeptide as claimed in any one of claims 1 to 4 which is provided as part of a fusion polypeptide.

6. An isolated or recombinant nucleic acid molecule which:

25 a) comprises or consists of the DNA sequences shown in Figure 1 (SEQ ID NO: 2) or its RNA equivalent;  
b) a sequence which is complementary to the sequences of a);  
c) a sequence which codes for the same polypeptide as the sequences of a) or b);  
d) a sequence which shows substantial identity with any of those of a), b) and c);  
or  
30 e) a sequence which codes for a derivative or fragment of the polypeptide shown in Figure 1 (SEQ ID NO: 2).

7. A vector comprising one or more nucleic acid molecules as defined in claim 6.

35 8. A host cell transformed/transfected with a vector as defined in claim 7.

9. A method of screening for and/or diagnosis of breast cancer in a subject which comprises the step of detecting and/or quantifying the amount of a polypeptide as defined in any one of claims 1 to 4 in a biological sample obtained from said subject.

40

10. An antibody, which binds specifically to a polypeptide as defined in any one of claims 1 to 4.

11. An antibody as claimed in claim 10, which is monoclonal, bispecific, chimeric, humanised or conjugated to a therapeutic moiety.
- 5 12. An antibody as claimed in claim 11 wherein the therapeutic moiety is selected from a second antibody or a fragment or derivative thereof, a cytotoxic agent or a cytokine.
- 10 13. A pharmaceutical formulation comprising at least one polypeptide as defined in any one of claims 1 to 4, at least one nucleic acid molecule as defined in claim 6 or at least one antibody as defined in any one of claims 10 to 12, optionally together with one or more pharmaceutically acceptable excipients, carriers or diluents.
14. A pharmaceutical formulation as claimed in claim 13, which is for use as a vaccine.
- 15 15. A pharmaceutical formulation as claimed in claim 13, which comprises one or more suitable adjuvants.
- 20 16. A method for the prophylaxis and/or treatment of breast cancer in a subject, which comprises administering to said subject a therapeutically effective amount of at least one polypeptide as defined in any one of claims 1 to 4 or fragment thereof, at least one nucleic acid molecule as defined in claim 6, or at least one antibody as defined in any one of claims 10 to 12.
- 25 17. At least one polypeptide as defined in any one of claims 1 to 4, at least one nucleic acid molecule as defined in claim 6 or at least one antibody as defined in any one of claims 10 to 12 for use in medicine.
- 30 18. The use of at least one polypeptide as defined in any one of claims 1 to 4, at least one nucleic acid molecule as defined in claim 6 or at least one antibody as defined in any one of claims 10 to 12 in the preparation of a medicament for use in the prophylaxis and/or treatment of breast cancer.
19. The use as claimed in claim 18 wherein the medicament is a vaccine.
- 35 20. A method of screening for compounds that modulate, i.e. up-regulate or down-regulate, the expression of a polypeptide as defined in any one of claims 1 to 4, which comprises the step of determining the presence or absence and/or quantifying at least one polypeptide as defined in any one of claims 1 to 4 or at least one nucleic acid molecule as defined in claim 6 in a biological sample.
- 40 21. A method for monitoring/assessing breast cancer treatment in a patient, which comprises the step of determining the presence or absence and/or quantifying at least one polypeptide as defined in any one of claims 1 to 4, at least one nucleic acid molecule as defined in claim 6, or at least one antibody as defined in any one of claims 10 to 12 in a biological sample obtained from said patient.

22. A method for the identification of metastatic breast cancer cells in a biological sample obtained from a subject, which comprises the step of determining the presence or absence and/or quantifying at least one polypeptide as defined in any one of claims 1 to 4, at least one 5 nucleic acid molecule as defined in claim 6, or at least one antibody as defined in any one of claims 10 to 12.

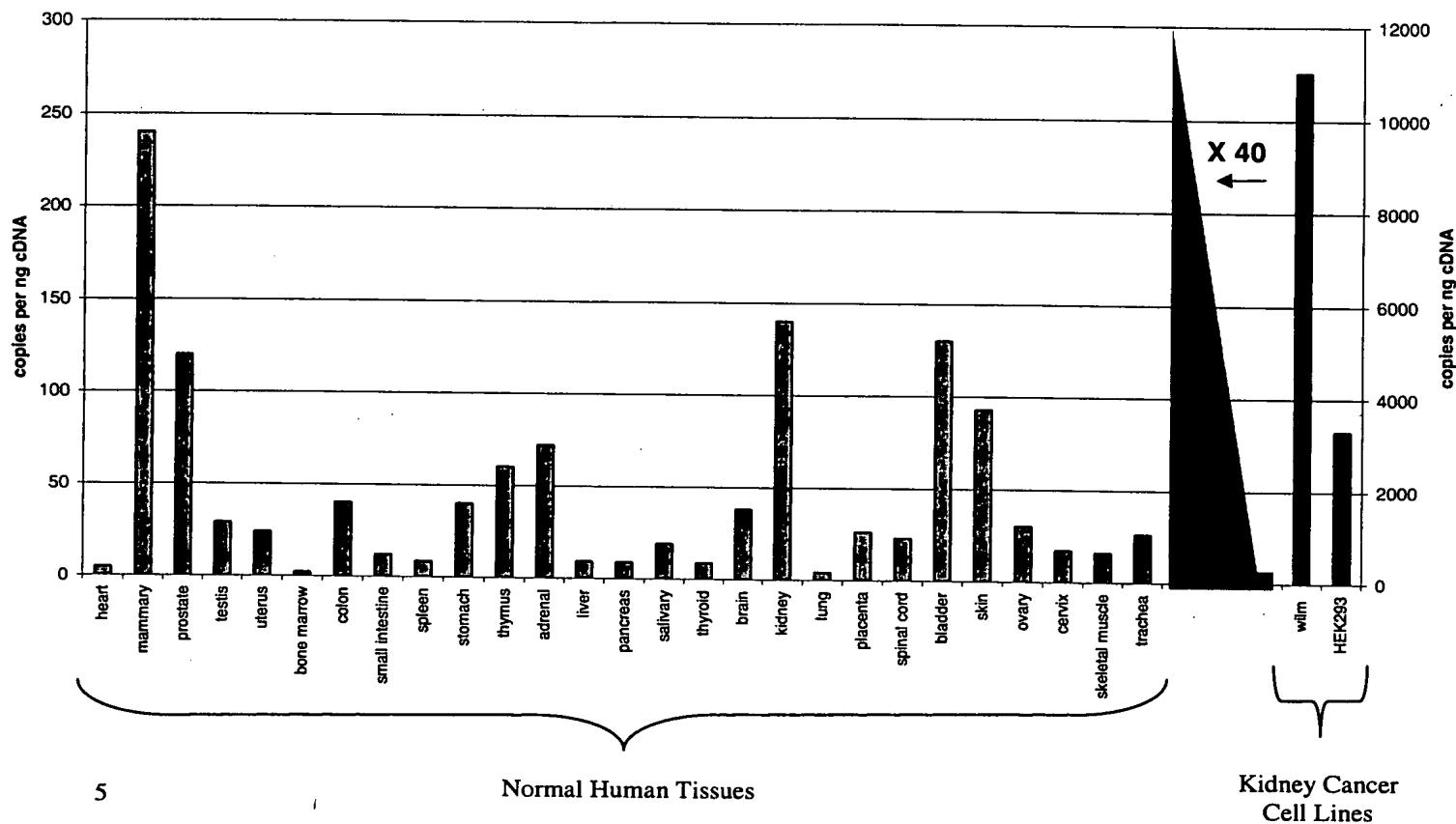
**Abstract****Protein**

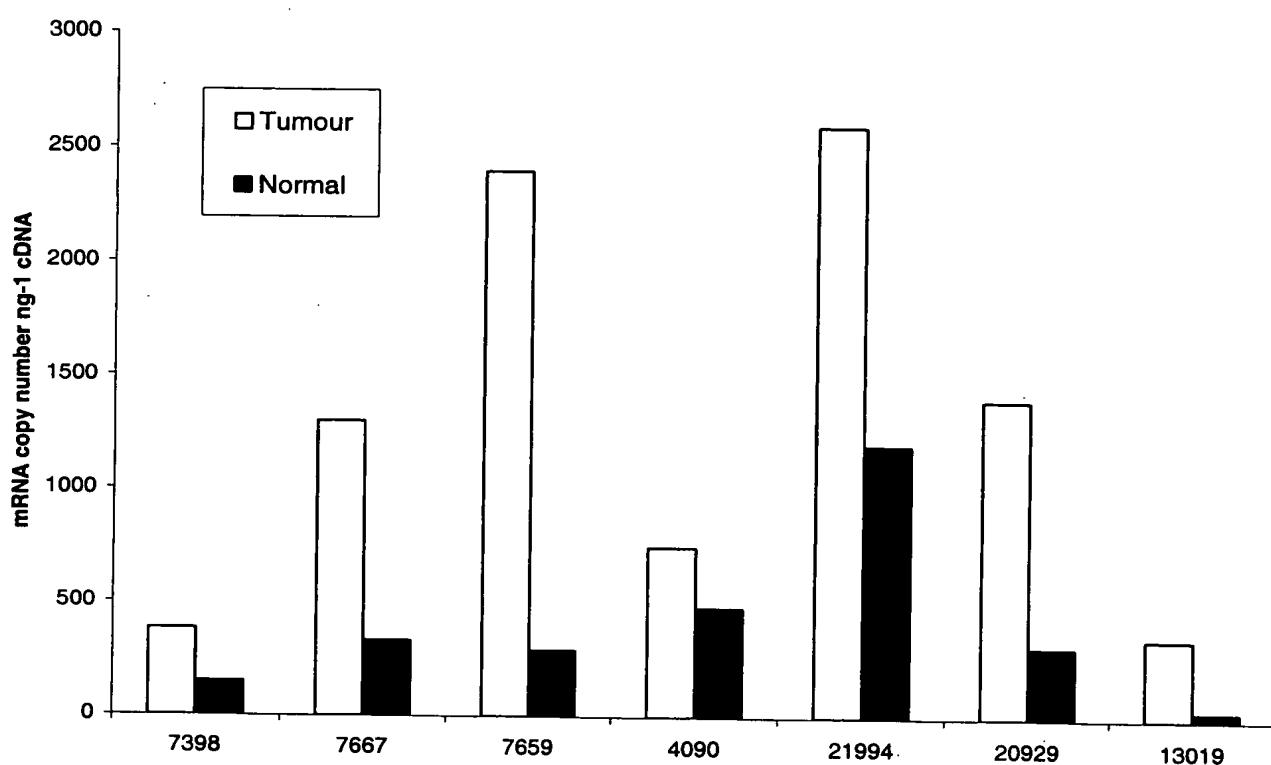
5 The present invention provides a protein (BCMP 101) isolated from breast cancer cell line membrane preparations, compositions comprising the protein, including vaccines and antibodies which are immunospecific for the protein. The use of the protein in the diagnosis, screening, treatment and prophylaxis of breast cancer is also provided.

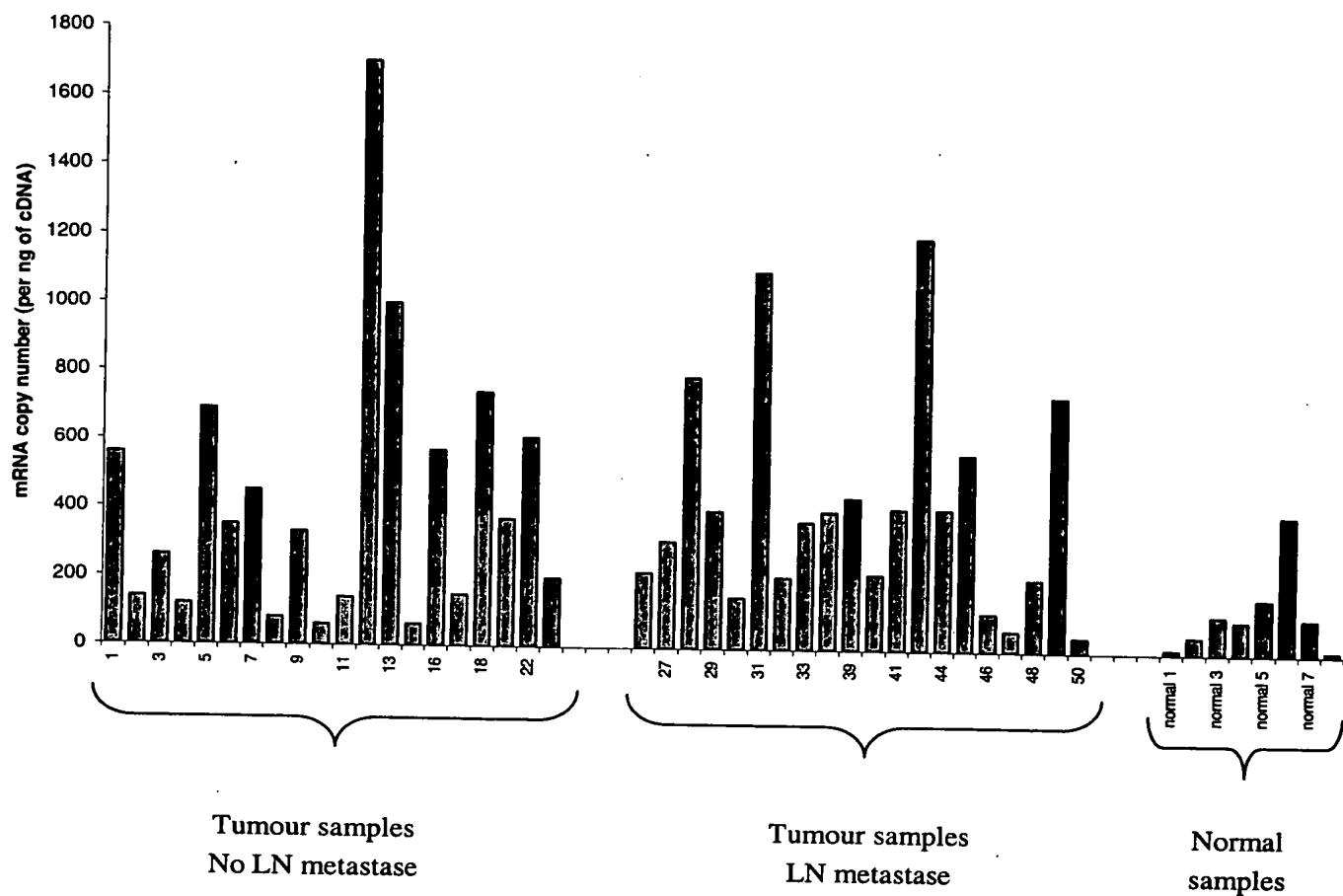
Figure 1

5	TGTGCAAATGACCCCTGGAGTTGGTTCGCTTTCTCCCTTGGGGCGGTGTGAACGTGTGT CCGCAGCGTG <b>AT</b> GGCAACCAGGTGGAGAAATTGACCCACCTAACGGAAAGTCC M G N Q V E K <u>L</u> T H L S Y K E V P	60 120 17
10	CACGGCCGACCCGACTGGCGTGGACGGGACGACGGGCCCCCATGGGGTCTCCTACAT <u>T</u> A D P T G V D R D D G P R I G V S Y I	180 37
15	TTTCTCCAATGACGATGAGGACGTGGAGGCCGCAGCCGCCCTCAGGGCCAGATGGCGG F S N D D E D V E P Q P P P Q G P D G G	240 57
20	CGGCTTGGCGACGGTGGGACGGGCCGCCGCCAGCCGAGCCCTACGATCCCG G L P D G G D G P P P Q P Q P Y D P R	300 77
25	GCTGCACGAGGTGGAATGCTCCGTGTTCTACCGGGACGAATGCATCTACCAAGAAGAGCTT L H E V E C S V F Y R D E C I Y Q K S F	360 97
30	CGCGCCGGGCTCGCGGCCGCTGAGTACCTACACGCCCGAGAACCTGCTCAACAAGTGC A P G S A A L S T Y T P E N L L N K C K	420 117
35	GCCGGCGATCTGGTGGAGTCGTGTCGCAGGCTCAGTACCCGCACTGGCCGTATATGT P G D L V E F V S Q A Q Y P H W A V Y V	480 137
40	GGGTAACCTCCAGGTGGTGCACCTGCACCGGCTGGAGGTGATTAACAGCTTCCCTGACTGA G N F Q V V H L R <u>L</u> E V I N S F L T D	540 157
45	CGCCAGCCAGGGCGTCGGCGCGTGGTCAACGATCTGTACCGCTACAAGCCGCTAAG <u>A</u> S Q G R R G R V V N D L Y R Y K P L S	600 177
50	CTCCAGCGCCGGTGGTGCACCGCGCTGGCGCACGTGGTGCAAGGAGCGCGAGCTGAG S S A V V R N A L A H V G A K E R E L S	660 197
55	CTGGCGCAACTCGGAGAGTTCGCCGCTGGTGCCGCTACGGCAAGCGCGAGTTCAAGAT W R N S E S F A A W C R Y G K R E F K I	720 217
60	CGGCGGCAGCTCGGCATCGCAAGCAGCCCTACCGGCTGCAGATTCACTGTCGGCGCA G G E L R I G K Q P Y R L Q I Q L S A Q	780 237
65	GCGCAGCCACACGCTCGAGTCCAGAGTCTAGAGGACCTGATCATGGAGAAGCGACGCAA R S H T L E F Q S L E D L I M E K R R N	840 257
70	CGACCAGATCGGGCGCGCGGCCGTGCTGCAGGAGCTCGCCACCGCACCTGCACCCGGCGGA D Q I G R A A V L Q E L A T H L H P A E	900 277
75	GCCGGAGGAGGGCGACAGCAACGTGGCGGGACTACGCCGCCCTCCGGGCGCCCCCTGC P E E G D S N V A R T T P P P G R P P A	960 297
80	GCCCAGCTCCGAGGGAGGAGGACGGAGAGGAGGGCAGTGGCACACT <b>G</b> ATGGCGAGCTGAGCGCA P S S E E D G E A V A H *	1020 310
85	GAGCTGCGAAGGGAACTGT'TTCAGTAGCAGCC	1054

Figure 2



**Figure 3**

**Figure 4**

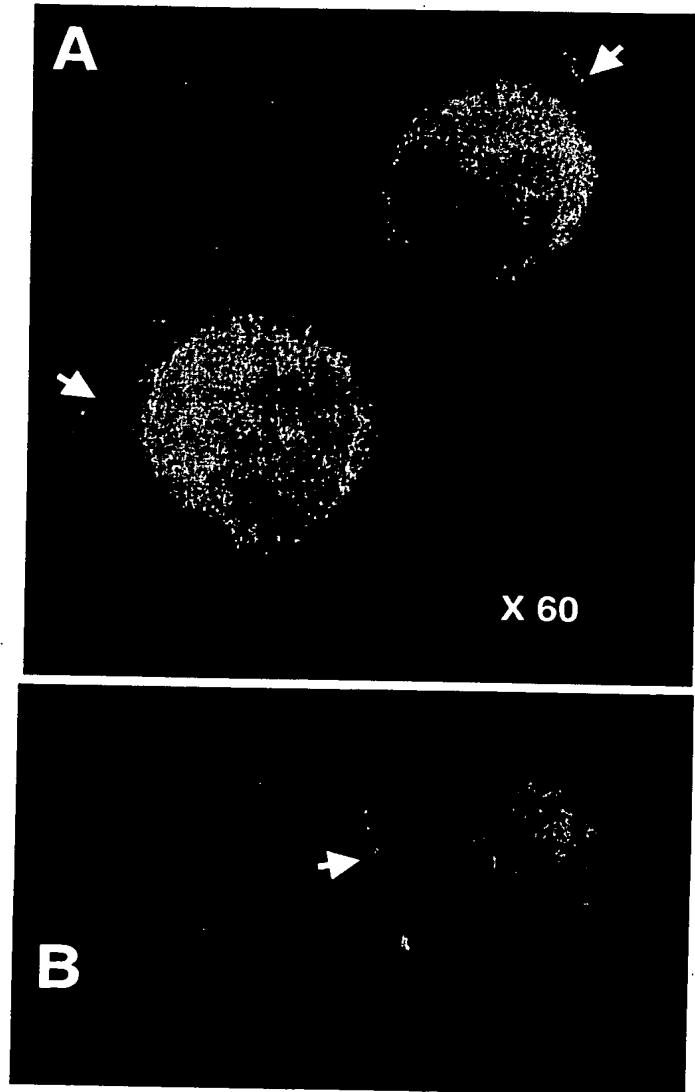
**Control**

**BCMP 101**



**Figure 5**

**Figure 6**



**Figure 7**

